Using Continuous Improvement Approaches to Achieve Scale: Implications for Depth, Spread, Ownership, and Sustainability

Marisa Cannata, *Vanderbilt University*

Stacey A. Rutledge, *Florida State University*

Christopher Redding, *Vanderbilt University*

Thomas M. Smith, *University of California-Riverside*

Mollie Rubin, *Vanderbilt University*

This paper is being presented at the annual meeting of the American Educational Research Association. April 27-May 1, 2017, San Antonio, TX.

Acknowledgment: The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305E100030. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Abstract

Recent history of school reform has numerous examples of promising programs that are developed and shown to have a positive impact on student outcomes, but encounter significant challenges regarding scale and sustainability. These challenges include building teacher buy-in, attending to the organizational context in which innovations are enacted, and aligning new practices with the broader array of programs that exist in schools. This paper is situated within the broader context of identifying how to more effectively scale and sustain effective practices. By both enacting a new approach to scale and studying how key principles of the approach contribute to scale, this paper provides informs how improvement science and design-based implementation research can lead to meaningful changes in schools. We find that an improvement model that emphasizes co-construction of practices in with practitioners, continuous improvement, and an authentic partnership between researchers, developers, and practitioners contributed to building internal ownership of the innovations, but encountered dilemmas over depth, spread, and sustainability.

Using Continuous Improvement Approaches to Achieve Scale: Implications for Depth, Spread, Ownership, and Sustainability

School reform has numerous examples of programs that have a positive impact on student outcomes and demonstrate initial implementation success, but encounter significant challenges while scaling and sustaining the program (Datnow, 2005; Glennan, Bodilly, Galegher, & Kerr, 2004). Many of these problems arise due to lack of teacher or organizational capacity to enact the reform (Desimone, 2002; Spillane, Reiser, & Reimer, 2002). Other challenges stem from the fact that local stakeholders may not consider programs demonstrated to be effective in other places relevant for their context (Coburn & Turner, 2011; Fishman, Penuel, Allen, & Cheng, 2013). Some reforms also fail to achieve scale because there is not sufficient infrastructure to support deep and sustained engagement (Hopkins & Woulfin, 2015; Peurach & Neumerski, 2015). Overall, achieving success at scale requires grappling with puzzles involving the context of local schools and districts, the expectations of the reform itself, and capacities of the organization scaling the reform (Cohen, Peurach, Glazer, Gates, & Goldin, 2013).

To overcome these challenges, researchers have called for new approaches to achieve educational improvement at scale that emphasize systems thinking, change ideas designed in the context in which they will be implemented, and iterations of testing, modifying, and improving the changes until improvement goals are achieved. Two approaches currently being advocated to address these challenges include improvement science and design-based implementation research (Bryk, Gomez, Grunow, & LeMahieu, 2015; Cohen-Vogel et al., 2015; Fishman, Penuel, Allen, Cheng, & Sabelli, 2013). While these approaches have multiple forms of organization and interaction, they share a focus on a shared problem of practice, the development

of a shared set of theoretical and empirical understandings of that problem, and the use of methodological approaches that allow for iteration and improvement. These approaches are made possible by involving researchers and practitioners working in partnership to design, implement, and scale educational innovations (Coburn, Penuel, & Geil, 2013). There is evidence that innovations that are developed through research-practice partnerships (RPPs) have positive outcomes for students (Sowers & Yamada, 2015). Less is known about how RPPs can support the spread and scale of the innovation (Coburn & Penuel, 2016).

The development of these new approaches to achieving scale has occurred along with a shift in scholarship on scaling up educational innovations (Cannata & Rutledge, in press). Coburn (2003) brought needed theoretical understanding to the definition of scale by arguing that scale is not only about the numbers of educators involved, but whether the reform results in deep and sustained change in educator behavior, beliefs, and norms. Increasingly, scholars and program developers recognize that "the ultimate goal of scaling up is sustainable educational improvement rather than to merely expand the use of a given educational innovation" (Sabelli & Harris, 2015, p. 14). With this shift comes the realization that scaling up an innovation requires a focus not only on concrete practices, but the theory of change behind the practices (Bradach, 2003; Elmore, 2016). Elmore reflects on his decades of work trying to scale educational reforms by noting, "'scale' for its own sake is less important than demonstrating that powerful ideas work in diverse environments" (Elmore, 2016, p. 533).

A related shift in approaches to scale up is a greater focus on the *process* of scaling up. While Coburn's framework extends the domains of scale, there is still a focus on scale (deep and sustained improvement in educational practice) as an outcome to be achieved (Coburn, 2003; McDonald, Keesler, Kauffman, & Schneider, 2006). Dede (2006) extends Coburn's framework

to include evolution. This shift is further evident in research questions that move from "to what extent has a given program or practice achieved scale?" to "what supports, processes, and infrastructure contributed to the scaling of a particular program or practice?" For example, a special issue of the *Journal of Educational Change* focused on "Bringing Effective Instructional Practice to Scale" highlighted the processes established by six instructional reforms that were successfully scaled up. Across these case studies of improvement at scale, there is an emphasis on a deep change in the culture of learning, local ownership of the learning agenda, and a system of continuous improvement (Fullan, 2016). Other research on scaling up has focused on the need for systemwide infrastructure to support improvement (Hopkins & Woulfin, 2015; Peurach, 2016; Sabelli & Dede, 2013).

This paper is situated within this evolving approach to scaling up. We argue that, while understanding the process of scaling up is important, achieving improvement at scale requires attention to both the processes and outcomes of scaling up. The research described here comes from a six-year partnership between our research team and two large districts to scale effective high school practices. This partnership had three improvement principles that reflect the new approaches to achieving scale: co-designing in context, an emphasis on continuous improvement, and leveraging expertise a research-practice partnership. The research question focuses on: To what extent, and in what ways, did these improvement principles result in depth, spread, shift in ownership, and sustainability? After situating the principles of our approach to scaling up in the existing literature, we describe the context in which this research was conducted and the data and methods of analysis. We then provide evidence on the extent to which each principle was successful in achieving scale.

Framework

New Approach to Scale

With previous challenges to achieving scale in mind, we developed a model with three improvement principles designed to address the scale-up challenges (Cohen-Vogel, Cannata, Rutledge, & Socol, 2016). This model embeds elements of improvement science and design-based implementation research, as there is a focused problem of practice, collaborative design, and attention to the organizational context in which the improvement work is situated (Bryk et al., 2015; Fishman, Penuel, Allen, Cheng, et al., 2013). As such, our improvement model is part of this emerging family of new approaches to scale effective practices.

The first improvement principle is that the innovation is built on practices shown to be effective in the partnering district and co-constructed with local educators. While a traditional approach to implementation considers adaptation a problem of low fidelity to the design (Durlak & Dupre, 2008), our model allows for principled adaptation as practitioners develop a sufficiently a deep understanding of the design philosophy and program components to productively adapt components with integrity to the underlying goals (Cohen-Vogel et al., 2016; Redding, Cannata, & Taylor Haynes, in press). This approach is consistent with educational implementation research, which has long noted mutual adaptation can help local stakeholders focus on their unique needs and learn from implementation efforts (Datnow & Park, 2009; McLaughlin, 1976; Siskin, 2016). Further, emerging scale-up approaches have a greater emphasis on integrity of implementation and adaptive integration, rather than fidelity to the original design. That is, re-focusing on educational improvement at scale means recognizing that improvement comes from integrating new practices with existing systems, as educators take the

effective practices, or standard work process, that are being scale and "integrate a standard work process into new contexts" (Hannan, Russell, Takahashi, & Park, 2015, p. 496).

Historically, reforms that achieved the most scale in terms of number of classrooms are those that practitioners feel are very adaptable (Cuban, 1998). Literature on scaling in business also recognizes the power of adaptations, as innovations can be used in novel ways and the process of improvement is never finished (Furr & Ahlstrom, 2011; Sutton & Rao, 2014). Yet adaptations can succeed or fail, with the innovations themselves evolving as the designers revise the theory of change as they observe how the innovation is adapted into specific contexts (Dede, 2006). As scaling up shifts from the transfer of a specific program to the adaptive integration of effective practices into new contexts (Hannan et al., 2015), there is a greater need for improvement efforts to help practitioners understand not only the innovation itself, but the theory behind the innovation (Thompson & Wiliam, 2008). By combining the "know-how" with the "know-why", practitioners can adapt reforms in ways that stay true to the underlying theory of change.

The second improvement principle is a continuous improvement model that uses rapid Plan-Do-Study-Act (PDSA) cycles to generate knowledge about the innovation, its implementation, and its outcomes (Bryk et al., 2015; Cohen-Vogel et al., 2015). This process contributes to achieving scale as we focus not on simply monitoring implementation, but also on building shared knowledge about the interaction between the innovation and the contexts in which it is implemented (Means & Penuel, 2005). The continuous improvement testing also focuses the local adaptation as each test provides knowledge about the implementation of specific adaptations in unique contexts to make judgments about those adaptations (Langley, 2009).

Navigating this uncertainty and complexity requires continuous learning among practitioners and those developing the innovation (Peurach, 2016). Practitioners need to try out new reforms, collect evidence on their effectiveness in that context, and iterate (Hannan et al., 2015; LeMahieu, Grunow, Baker, Nordstrum, & Gomez, 2017). By collecting and analyzing context-specific evidence, partnerships can improve both the innovation practices and the system that supports their implementation (Penuel, Fishman, Cheng, & Sabelli, 2011).

The third improvement principle is working in a research-practice partnership that fosters local leadership (Bryk et al., 2013; Cannata, Cohen-Vogel, & Sorum, in press; Coburn et al., 2013). By engaging teachers and leaders in design and development, we aim to build buy-in for the innovation and take advantage of local expertise about the context, while simultaneously building local knowledge about the innovation (Lewis, 2015). Elmore (2016) suggests that since education is a people-oriented profession, large-scale educational improvement comes from establishing powerful learning communities that engage around central ideas of practice. Indeed, all forms of new approaches to scale include some type of network or partnership (Coburn & Penuel, 2016; Fishman, Penuel, Allen, & Cheng, 2013; LeMahieu et al., 2017; Peurach, 2011; Peurach & Glazer, 2012).

One reason that networks and partnerships are important to achieving scale is that they can leverage multiple forms of expertise. The problems that face our educational system are deep and complex, requiring diverse perspectives and types of expertise, to see whole systems (Bradach, 2003; Bryk et al., 2015). Partnerships allow individuals and organizations with different forms of expertise to align their efforts and increase the likelihood of success (Bryk, Gomez, Grunow, & Hallinan, 2011). Further, organizational learning that facilitates adaptive integration requires continuously engaging multiple stakeholders in a process of ongoing

feedback that engages both top-down and bottom-up ways of interacting (Chambers, Glasgow, & Stange, 2013; Fullan, 2016).

Defining Scale: Scaling In and Scaling Out

To determine whether our model supports improvement at scale, we need a framework that allows us to assess the extent to which scale is being achieved. Scale, in Coburn's (2003) framework, includes spread, depth, shift in reform ownership, and spread. We adopt this definition of scale, although we recognize that the innovation itself can evolve as local adaptations inform the developer's thinking (Dede, 2006). Using this definition, we define *scaling-up* as involving both *scaling-in* to individual schools and *scaling out* throughout a district. *Scaling-in* focuses on processes inside individual schools, recognizing that part of achieving scale is enacting deep change in the beliefs, norms, and practices in schools (Redding et al., in press). In assessing the extent to which the key principles have achieved scale within schools, we attend to issues of ownership, beliefs, capacity of individuals in schools, and the school organizational processes to support sustainability.

Scaling up within large districts requires attention beyond individual school stakeholders. We define *scaling-out* as the process of spreading to new schools, while also establishing ownership and mechanisms to support depth and sustainability in central office processes. For example, district-wide scale requires district leadership and key organizational personnel to have ownership over the initiative and commitment to its success, establish structures to help schools deepen their implementation, reallocate resources, establish communication channels, and institute other mechanisms for sustainability. Taken together, investigating scale up efforts in a large district should focus efforts at both the school and district levels.

Context

This study is situated in partnerships between research universities, a curriculum and program developer, and two large, urban school districts. While the focus of the work in each district is unique, the improvement model was similar across both districts. The work began by researchers intensively studying high schools in each district, with the goal of identifying programs, processes, and practices that distinguished higher from lower value added high schools and then to use these findings to inform an iterative process of design, testing, revision, and implementation of an innovation improve student outcomes. These designed innovations in both focused on student-centered practices to develop socioemotional skills in students, although there were specific areas in which the innovations were distinct (Rutledge & Cannata, 2016). In District A, the innovation was known as Student Ownership and Responsibility (SOAR) and in District B, the innovation was known as Personalization for Academic and Social Learning (PASL).

As shown in Figure 1, the improvement model was implemented in each district over four school years and included four main phases. The first phase, design, involved the design of the innovations by a District Innovation Design Team (DIDT) that was comprised of about 20 people that included teachers, school-based administrators, central office leaders, program developers, and researchers. The teachers and school-based administrators were from three high schools that were designated as the first sites of implementation (known as innovation schools) and other high schools in the district that might later implement the innovation. In the development phase, School Innovation Design Teams (SIDTs) were formed in each innovation school, that included their representatives on the DIDT (known as school coordinators) and 6-8 other teachers or support staff. The SIDTs were tasked with further developing and piloting the

innovation that emerged from the design phase. In the implementation phase, the SIDTs in the innovation schools led the implementation and engaging in continuous improvement., supported by the developers. In the scale out phase, four to five additional high schools created SIDTs and began phasing in implementation in their schools. Across these four years, DIDT and SIDT meetings offered opportunities for capacity building, collaborative learning, and organizational planning that were facilitated by the developers and researchers. Such meetings were initially monthly for two full days and later became quarterly full-day meetings.

While the overall improvement model was similar across the districts, there were a few notable exceptions. First, the work in District B began six months before District A, and two key challenges emerged that were incorporated into District A when the work began. One challenge was the tendency for District B DIDT members to interact primarily with their campus-based team and to focus more on designing for improvement in their schools than designing for improvement in schools across the district. In District A, therefore, stronger efforts were made by developers to emphasize the district-wide focus of the DIDT's work and, for many activities, assign members to cross-school groups. The other distinction was that the initial innovation design in District A needed to be further specified prior to implementation than was the case in District B. While this specificity was developed over time in District B, the DIDT launched in District A with a more detailed design and supporting materials developed. The development phase also showed differences across the districts. In Fall 2013, the program developers and researchers in both districts were concerned that the design process had not achieved enough specificity for schools to plan for implementation. This led to a "repositioning" that had the developers take a more active role in developing specific innovation practices or tools. While this repositioning occurred in both districts, it was emphasized more in District B as they were

deeper into the work. Also, while the PDSA process was introduced similarly in both districts, in District B there was an emphasis on having all innovation schools test the same PASL practice, while that assumption was not present in District A. The approach to the implementation and scale out phases were similar across the districts, although contextual differences between districts led to some differences. In particular, an extended superintendent search in District A and resulting transition in district leadership led to less engagement by senior district leaders in the DIDT.

Data/Methods:

Data for the analyses reported in this paper come from three sources in these districts: (1) observations and artifact collection from the DIDT/SIDT meetings, (2) interviews and focus groups with teachers, administrators, SIDT members, and students from school fieldwork visits, and (3) additional interviews with DIDT/SIDT members at several points in time.

First, DIDT/SIDT meetings offer fruitful data for exploring the improvement process, as these were working meetings where participants shared data and discussed what they were learning through the process. At each DIDT/SIDT meeting, researchers completed structured field note logs that captured the nature of the activity in participants were engaging, the individuals involved, content of the discussion, and the level of engagement. Audio recordings provided additional data on small groups that a data collector could not observe. These recordings were partially transcribed and analyzed through a reflection form organized according to the coding scheme described below. In addition, participants completed feedback forms and the research team wrote a summary reflection on the day's events. Finally, materials distributed or created during the meeting were collected. Across the four years of meetings, we recorded 925

hours of meeting audio and collected 35 sets of feedback forms, 43 researcher reflection forms, 118 fieldnote logs, and 988 artifacts.

Second, we visited each innovation and scale out school. Innovation schools were visited once prior to implementation to gather baseline information about the capacity and will to engage in this work, once or twice in the first year of implementation, and once in the second year of implementation. In 2015-16, we visited each innovation and scale out school once. In each visit, we interviewed SIDT members, teachers, administrators, and other school personnel, and conducted student focus groups.

Third, we conducted additional phone interviews with DIDT/SIDT members at several points in time, beginning just after the design phase until after our partnership had officially ended. Table 1 summarizes the fieldwork data collected in each district.

The network meeting and interview data help to describe the organizational features of the network and the types of information that spread throughout the network. Transcribed interviews and meeting data were analyzed as part of the project's framework for innovation design and development. This framework consisted of several *a priori* codes in addition to codes that emerged inductively from the data. The data were coded using a framework for design, development, and implementation that focused on attitudes and engagement in the improvement process; content of learning activities in meetings; extent to which the improvement process was collaborative, aligned with school and district priorities, and grounded in the design challenge; engagement in PDSA; role of DIDT/SIDT members; network leadership; facilitating conditions for implementation; implementation supports, implementation quality, and student experiences with implementation. This framework was mapped onto Coburn's (2003) definition of scale as follows: depth is reflected in changing beliefs, greater integrity when enacting practices, and

greater student responsiveness; shift in ownership is reflected in increasing will, more productive SIDT dynamics, and continuing engagement in continuous improvement processes; sustainability is reflected in increased capacity, greater alignment of local practices with innovation, greater engagement of SIDT members, allocation of local resources, and continued provision of support to teachers; and spread is reflected in greater frequency when enacting practices and a broader program reach. The coding process was iterative in nature with members of the research team comparing coding to ensure a consistent understanding and application of codes (Corbin & Strauss, 2008).

After the data were coded, analytic memos were prepared for each district by major code area, summarizing the themes that emerged in those codes for each district in each phase of the work and including evidence (i.e., quotes, vignettes of meeting activities) that substantiated those themes. Then, researchers systematically went through these memos searching for evidence about how each improvement principle contributed to each of the four components of scale.

Findings

Co-construction of practices in context

In this section, we provide evidence for the different ways that the co-construction worked to support scaling up in our RPP. We first discuss how co-construction in the different phases empowered practitioners to spearhead local improvement efforts. We then turn to the challenges of co-construction in the design process. We also discuss tension between ownership of the innovation, on the one hand, and depth of change, on the other hand.

Co-construction empowered practitioners. Overall, there is considerable evidence that our model's approach of basing the design and development work in the district context, and facilitating co-construction of practices at the school level, led to a sense of ownership among

those involved. Across all phases of the work, members repeated that their direct involvement in developing the PASL/SOAR innovation led to their buy-in and sense of ownership. For example, one District B DIDT member described what she thought was the most important thing about the work by saying,

> "This process has been, from the start, a conception of people that work in the school, that work in the district, that understand the needs and what's going on and so I hope that the new folks that are joining us realize that this isn't just another one of those binder workshops where they're going to train us and then send us back; you know this was developed by the people that work alongside you and understand what you go through at your school…. that's what I appreciate about this particular endeavor."

Likewise, after a presentation on the history of the project to members joining the network at one of the summer institutes, a District A DIDT member noted, "you could see people sit up in their chair when they heard it was developed by teachers in the district."

The nature of co-construction differed during each phase of the model. During the design phase, practitioners participated in co-construction by participating in monthly meetings aimed at turning the design challenge findings into an implementable reform. During this phase, we found that practitioners needed time to familiarize themselves with the idea that they were being asked to participate in adapting PASL and SOAR to different school contexts. Participants understood the logic, but they struggled with what that entailed. For example, a District B DIDT member commented, "We're supposed to be active participants in this but it felt like 'sit and get' on Tuesday. And a lot of that is just giving us the basics because we have no base knowledge yet." During the development and implementation phases, participants were introduced to and participated in multiple rounds of PDSA, which we will discuss in more detail below. During

these phases, practitioners engaged in co-construction when they worked as a team to identify the elements of the innovation they would work to implement. They also engaged in co-construction—although the level of this differed by school—when they shared the goals and findings with non-SIDT teachers who were implementing the innovation. In this way, co-construction occurred at three levels: the DIDT, the SIDT, and when the SIDT involved other teachers.

Challenges inherent in co-construction.

Despite the empowerment DIDT and SIDT members felt through the co-construction process, there were also inherent challenges and limitations. During the design Phase, there was tension between the developers and researchers, on the one hand, and practitioners, on the other, regarding the degree of freedom that practitioners had in adaptation. While the developers and researchers valued the voice of practitioners, they also grew worried that the practitioners were veering too far off the design principles of PASL and SOAR. Drawing from earlier feedback and the Year 1 findings, developers and researchers introduced "non-negotiables" of PASL and SOAR. Co-construction continued, but it was delimited by addressing these elements of each innovation.

We call this recognition of the frustration felt by DIDT/SIDT members, and concerns by the program developers and researchers that the development process was stalling, "repositioniong" of the work. With this repositioning, program developers began taking a more hands-on role in developing the innovation. Many members of the DIDT and SIDTs, for the most part, appeared to accept the "repositioning" without significant resistance. Some participants, in fact, expressed confusion regarding how, exactly, the model had fundamentally shifted – one SIDT member in District B, for instance, noted that s/he didn't "really know how

this is different from what we were doing". Others reacted with more concern. One SIDT member in District B, for example, explained that "my experience is that teachers are very willing to go the extra step, but they need to know where they're going. So if you say let's slow down to figure this out and then we'll let you know – it's going to be a little frustrating, and we'll lose a little credibility." Another SIDT member said, "it was very influenced -- quite prescribed. At a certain point in the beginning when I came on board, it almost seemed like we were building the program. But as time progressed, it appeared that it became more and more prescribed and this is what they wanted to essentially see done." Similarly, in District A, some members reacted to the respositioning with some resistance. Most members appreciated the specificity of the materials provided by the program developers and were highly engaged in providing feedback on them, however, the most significant concern raised by practitioners was that their voice, and the significant independent development work they had undertaken, would be lost with the advent of a more "top-down" model.

A similar challenge related to depth emerged as teams began preparing for implementation. While the DIDT and SIDT members felt ownership over the innovation due to their ability to develop in their context, they were concerned about how to build buy-in among their school more generally. A DIDT member from District B noted that PASL was "still a somewhat vague concept" for the other administrators and teachers at her school, because "they don't have the kind of details that they want to know." Later, another DIDT member expressed the same concern of building buy-in among others in her school, "my role on the campus is to be the answer – they come up to me and they want me to give them the answer […] I couldn't give them the details on this, and I think through this process, we have to be careful that we're not throwing out a teaser and then not having enough information to share with them. That creates a

little bit of mistrust, to be honest with you […] they don't trust if you can't give them answers to their questions."

This frustration was also felt in District A, with a DIDT member saying, "I think that it's necessary that we first finish up what the plan is. So as far as me being able to answer that question fully, 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." However, this frustration came from wanting the PASL/SOAR innovation to work rather than from not believing in the principle of designing in context. Members who pushed most in beginning became big proponents in the end. An SIDT member in District A commented at the end of the development phase: "…[E]ven though it may be a little frustrating just, you know, at times because of our busy schedules and what we're doing as teachers, and trying to make time for it, it's worth it."

There were differences between the two districts, though, that lead to tensions in co-construction. In District B, feedback from practitioners during the design phase lead developers and researchers to include administrative involvement at the SIDT level. School coordinators at the three innovation schools were also assistant principals (APs) at the school. In a top-down environment like District B, AP support signaled to teachers that this was something worth implementing; the negative, though, was that sometimes the AP took on too much of a role and did not distribute to the teachers or even cut them out.

In District A, the SIDT members were largely teachers. They took an active role at their schools in implementing SOAR, with members were so enthusiastic of the co-construction process that they spoke of trying to create a similarly collaborative and inclusive process for implementation in their schools. For example, one school described how they introduced a specific practice to their teachers and, in doing so, emphasized that teachers have freedom in

modifying their materials for their classroom. In describing this to the whole SIDT, one member commented, “we felt a breath of fresh air with our teachers, it’s like we have ownership from the teachers.” As another example, at the end of the first year of implementation, innovation school members were asked to give advice to the scale out schools. One member said, “I think it needs to be inclusive and inviting because if you’re exclusionary or teachers feel like they can’t join, they’re not going to participate in it.” Another member mentioned the importance of transparency from the SIDT to the rest of the school so they know what is going on.

At the same time, these teacher leaders worried that their authority might be compromised when other participants asserted a position. Participants went so far as to describe how the introduction of the lessons developed by the program developers hurt buy-in on their team, but this abated over time. As a SIDT member in District A explained,

> “I would say buy in was weakest in the fall and gained strength over the spring.…[the program developers] put together four lessons, two on growth mind set and two on problem solving. They were piloted by the schools, by the SIDT members, but basically they were piloted as is. …[we] had continual concerns about these particular lessons, and in particular, they were wondering about how are we actually going to convince teachers to give up four whole lessons to introduce these ideas, and in the spring, as they began working on an implementation plan, there was kind of a release of kind of -- it became clear that the school didn't need to adhere to the format of four lessons, and they began to come up with different ways to organize the initial delivery of this content. And once that happened, I think buy in increased a lot.(1333)

They also voiced concern that the lack of formal leadership impeded the spread of SOAR. When administrator support was not sustained or school priorities shifted, it challenged the sustainability of SOAR.

Co-construction and the dilemma of depth. Practitioners in both districts were more likely to support the innovation when they could identify policies, programs or practices already in place at their school and therefore similar to something they were already doing or already had experience doing. We observed numerous examples of this across both districts. In District B, for example, SIDT members from one school often pointed to programs offered by the media specialist—a SIDT member—as promoting PASL. Similarly, SIDT members at another school would identify a mentoring program initiated by the principal prior to the innovation, as evidence of PASL. In District A, similarly, SIDT members adopted practices that met immediate teacher needs, instead of choosing practices that might have been more challenging to implement. At both sites, developers and researchers worried that by turning to either prior practices or more easily implemented ones, SIDT members were not exerting effort to implement new components as practitioners gravitated towards less ambitious and easy-to-implement practices that may have less impact on students.

Taken together, we observed that the DIDT and SIDT process engendered co-construction and local ownership much as intended. While initially cautious, practitioners in both districts welcomed and embraced their active participation in the development and implementation of the innovation. While we did observe challenges and limitations to co-construction, we believe that it helped to establish strong belief in the innovations as well as strong support at both the school and district levels.

Continuous Improvement and Scale Up

Challenges with continuous improvement. Several challenge limited the extent to which SIDT members adopted the Plan, Do, Study, Act (PDSA) process and applied it to their improvement work. These issues were generally consistent across districts, and included negative attitudes towards the process itself, challenges with the mechanics of successfully implementing PDSA, and organizational constraints to its successful application. Although most members agreed with the general principle of reflective inquiry as part of their practice, when PDSA was introduced, members in both districts viewed it as "bookish," "prescriptive," and "bureaucratic." An SIDT member in District B summarized this attitude:

> I do understand that Plan Do Study Act is – has to be re-evaluated over and again just to make sure that we are fine tuning and that we are not adding more but that less is more, so that we have quality rather than quantity. But the whole process of the way it was done seems so needling, so laborious, and – and drawn out that I think part of the SIDT team became very frustrated with it because it was meeting after meeting we're looking at the same thing again.

In District A, SIDT members attributed their dislike of PDSA to rigidity it tried to impose on their improvement process. An SIDT member in District A remarked, "I definitely feel like we incorporate the things that we learn into what we do next. I think we do a pretty good job of that, but it's a very organic and kind of agile process as opposed to one that keeps a good track of where we stand and what we've accomplished."

When SIDTs attempted to engage in the PDSA, they often struggled with the mechanics of this continuous improvement process. One of the most common challenges related to measurement. Although SIDT members in each district were immersed in data, they struggled to identify the outcomes they sought to change and determine how best to measure them. A developer in District A reflected, "Teachers are certainly used to looking at data. They look at data all the time. But, I don't think that they're used to thinking about how you design what data you

want to collect." For instance, after one SIDT shared their PDSA plan during a district meeting, the developers pushed the group to be more specific in their "recommendations" and also think about "how you would measure" whether or not teachers are forming positive relationships with their students. When one SIDT member asserted that it is "common sense" and "good practice", the developer pushed back again by asking "how would you know?" This particular exchange resulted in the developer asserting that they should identify measurable indicators.

An additional challenge related to PDSA's iterative approach to school improvement. A researcher in District A reflected:

> They had a hard time thinking about testing something. They said, 'okay, we need to develop everything that we're doing, and then we can figure out what we want to test'. We were trying to say to them, 'well, how do you know you want to do something unless you know it works?' But they wanted to keep focusing on, let's put the whole plan together before we think about testing anything.

The SIDT at one school in District A even disregarded the iterative approach, transforming PDSA into a year-long process for tracking implementation success. An SIDT member at the school explained:

> For our school PDSA was more effective when we stopped trying to evaluate all these different things in a year and instead we switched to overarching themes that we wanted to track. That has helped us to develop more focus for PDSA and we now have longitudinal data that we can look at from year to year from PDSA. PDSA isn't my favorite, but switching approaches and moving from a discrete perspective to an overarching approach has been helpful for my school.

Finally, organizational constraints limited the adoption of PDSA. The primary constraint was time to analyze the data and identify areas of improvement. An SIDT member in District B said, "We collected all this data, but we don't have time to look at it. I think a lot of our PDSA cycles get stuck because we collect all this data, but we never have time to sit down and analyze it." An SIDT member in District A elaborated, noting not only time but the administrative support as inhibitors to PDSA:

> What are your indicators going to be and how are you going to measure them? Once you do that you need to have time to study the results...you always need more time, I'm sure everyone is going to say that. You need the personnel to analyze the data and the admin support, for the action you need an established vision of SOAR on your campus, the next step after that is continuing the process and doing another PDSA cycle to continue the work.

These various challenges and organizational constraints resulted in inconsistent use of PDSA across schools, particularly in District A. While one school actively engaged in PDSA, another school spent time planning and looking at data, but not necessarily using the iterative PDSA framework, and the final school did not successfully complete any PDSA cycles. In District B, there was less variability across schools in terms of application of PDSA. The more successful use of PDSA can be linked to greater overlap with what was tested across schools. While SIDTs conducted common PDSA cycles for their first two cycles, in District B, they adopted a "parallel development" strategy, where all SIDTs tested the same change idea, even when it was implemented in differently across innovation schools. That being said, for the first year of conducting PDSA, SIDTs across both districts found it difficult to successfully complete iterative cycles. Teams would often complete one or two elements of the PDSA cycle, but rarely the whole process.

Build ownership by gathering feedback. Over time, schools in both districts had greater success in the PDSA process, which occurred alongside a shifted in the goals of PDSA. Initially, the goal of PDSA was to collect evidence of positive student outcomes. For instance, SIDT members were initially trained to identify the student outcome they expected to improve as a result of their change idea and make a prediction of the degree of change that would need to take place for this change idea to be considered a success. SIDTs in both districts shifted to focus on teacher outputs and perceptual evidence rather than student outcomes as the primary source of evidence of success. In District B, data collection efforts often focused on the frequency with

which teachers were employed PASL practices. In District A, SIDTs drew more heavily on teacher and student feedback to gauge the success of the innovation.

The focus on teacher feedback was vital for creating broader ownership over the improvement process in two innovation schools in District A. At one school, SIDT members sought feedback from their colleagues to customize the innovation to meet the needs of their teachers. An SIDT member explains, "I think—I mean—going through the [PDSA] cycle we've changed the lessons, and—or not changed them, but refined the lessons to fit what our campus is going to be using for professional development, which then means we're changing it to fit our campus." Even without direct involvement in the reform process, the incorporation of teacher feedback into the innovation increased the staff's receptiveness to the innovation. A teacher in a focus group remarked, "I feel like if [the SIDT] didn't put in as much work as they did, we, people wouldn't be as receptive." At that school, SIDT members' enthusiasm and knowledge helped generate ownership over the innovation among teachers uninvolved with the improvement process. Although this approach created a high degree of ownership, another school achieved even broader ownership by involving nearly half of the staff in the development of curriculum materials related to the innovation. An SIDT member commented:

> [T]here would always be like the data collection part, and then as we were, you know, kind of looking over the data, that's when we would like come up with the next PDSA thing, year. We were always collective and looking at data and figuring out what our colleagues wanted, needed, and trying to base our next actions based on that.

Continuous improvement and the dilemma of spread. PDSA was helpful in building ownership over the innovation. Yet, the process by which schools built ownership involved customizing the innovation to each local school context risked departing from a common district innovation that could be scaled to new schools. As one DIDT member in District A explained, PDSA was used to develop the innovation at the campus-level: "from the very beginning, we

have worked with the PDSA cycle as it was brought to life at each of the campuses." In other words, the initial use of PDSA was ambivalent as to how the innovation would be spread to new schools in the district. Yet, we found this internally focused continuous improvement process was not an inhibitor to spread. Tools that were developed at the innovation schools were adopted by other innovation schools and the scale out schools. Each district's approach to spreading the innovation was distinct, offering two models of how locally developed practices could be spread to new schools within the district.

In District B, although there was still an openness to adaptation to local context, the DIDT identified a set of core practices that would be implemented across schools. These practices included educator teams that follow students for three years, the assignment of teacher mentors for every ninth grade student, and the explicit teaching of a PASL curriculum. With a core set of practices, differences across campus were in terms of the grade levels in which PASL was implemented and the frequency with which these practices were implemented. Given some of these differences in how PASL was implemented, discussions during quarterly meetings centered on how they would make the innovation "more tractable" to spread beyond the innovation schools. Developers guided discussions that answered the question of what it meant to successfully implement PASL:

> We want to get to the time where we're having a really good conversation about what really works. When we're talking to someone else – how will we say, 'if you're really going to do PASL, this has high leverage, so you should do this' […] from a high level, we want to get to a point where we can identify what is truly important to do.... That's what PDSA process is all about – going through that improvement process.

In addition to the collection of practices, a district central office staff member emphasized that the expertise developed on SIDTs in the innovation schools would enable the spread of the innovation: "We're going to double, add on another 7 to get to 15. These cycles are going to

allow us to get those lessons learned…so you guys are becoming those resident experts…it's a way of thinking through these things in smaller chunks."

In District A, while there was some overlap in the practices that were adopted across schools, there was greater adaptation of the innovation to each school context. This greater openness to adaptation yielded greater diversity in the practices developed across the innovation schools. This approach gave schools the flexibility to create distinct practices, test their success using PDSA, and revise the practice before sharing it with other innovation and scale out schools. As each school developed and refined practices that would achieve the goals of the innovation, ongoing opportunities for sharing their learning with one another led to the identification of common practices that could be deployed to achieve the goals of the innovation, namely, increasing student ownership and responsibility. This approach yielded a set of practices that scale out and other innovation schools could choose from when they were defining what SOAR would like on their campus. One SIDT member from a scale out school remarked:

> What did you gain from being involved with other schools across the district? -- I've been able to take the meat without the bones. Everything that other groups or schools would share I was able to take it and bring it to my school, and give it to the team to modify and revise based on the needs of our students and the community. They may not have realized it came from this team, but everything we've done on a district level has helped with the work we've done at the school. This is unique because it comes from FWISD and teachers, that's a reason for us to be proud.

Another SIDT described this set of practices as "a bank of exemplar models, lessons, professional development opportunities that would support this kind of work, so that if a campus were to start up this initiative, or was hitting a couple of walls, they would have something to look to."

In a similar way that PDSA helped to spread reform knowledge among schools in the district, over time, it also became a hub for formalizing institutional knowledge on the DIDT.

Across both districts, PDSA gradually became a tool by which teams could formalize their learning in a way that improvement efforts could be sustained. During the 2015-2016 school year, a district central office staff member involved in the improvement work in District B emphasizes how they had built a culture "where this is a safe place to say hey, if you're struggling in a particular area, there's folks around that have figured some things out that can get better at what we're doing. These cycles are about what can we do to get better." A developer added: "That process is going to build on itself so that everybody is on board. This will help onboard the new schools as well as to continue to improve what you're doing currently." Similarly, the SIDTs in District A had gained increased comfort in using PDSA was the school and district-wide improvement efforts. An SIDT member said: "PDSA lets me take data and make it useful for others. We can present to people outside of our team and show the why and what of the work we're doing." This feature of PDSA as the hub of the network was emphasized in quarterly meetings in each district. A developer described PDSA as forming the basis of the DIDT's knowledge of what SOAR looks like across district. One of the ways the DIDT were described as being able to communicate their learning was at a district as whole, drawing data from multiple schools that are "shared across the district."

Research-Practice Partnership

Overall, there is mixed evidence about the extent to which the specific type of research-practice partnership used in this improvement initiative contributed to scaling PASL and SOAR in these districts. While there is some evidence that the partnership helped to increase local ownership, there were also tensions around creating a coherent districtwide identity for the work, confusion around the role of the DIDT, and a dilemma around who should "own" the work. These tensions illustrate how uncertain ownership and clarity about the various roles and

processes to support the work can inhibit developing a sustainable infrastructure for improvement. We describe each of these findings.

How the partnership contributed to ownership.

In both districts, the work began by establishing a districtwide team to develop the core components of the innovation. The goal was to create a single innovation that would be adapted to the needs of specific campuses. There were some successes in establishing this collaborative partnership and creating a sense of excitement within this team. For example, throughout the meetings, there were instances of laughter and joking among DIDT members. This good rapport appears to have engendered positive interpersonal relationships between members of the DIDT and frequently engaged in social conversation within and between activities. One DIDT member in District A linked the diversity of perspectives in the partnership to what he/she found rewarding about the process,

> I think it has been very interesting to see how kind of this group of people from across the district, you know, on the DIDT kind of comes together to try to grapple with some of these ideas … So I think it's been rewarding kind of seeing the potential for these kind of cross-role collaborations to be fruitful.

A DIDT member from District B similarly described the positive sense of collaboration among different partners: "it's been a real dynamic group. … I mean, it's been a good group of people working together from a practitioner standpoint, from a researcher standpoint, and then from a program developer, …We have people there where it doesn't become the next thing." [DIDT016].

Further, as the partnership progressed and the SIDTs joined the work, there was evidence that some DIDT and SIDT members took ownership of the improvement process. For example, DIDT members in both districts were involved in describing the innovation to new SIDT members, which demonstrated that this was not an external innovation, but one developed by the

DIDT. Further, DIDT members in both districts increasingly took leadership roles in facilitating activities at network meetings and the summer institutes, as part of the "gradual transfer of leadership" from the program developers to the district over time, particularly in the implementation and scale out. In June 2015, for example, both districts held summer institutes, which were the first opportunity to introduce the scale out schools to PASL/SOAR. In District B, DIDT members led the introduction of the PASL core components to the scale out schools through a structured poster session. In District A, a district administrator (and DIDT member) led a clinic labeled "SOAR 101" for the scale out schools and each innovation led a workshop on what they have learned through implementing SOAR. In 2015-16, district leaders in both sessions were often leading network meetings and calling on DIDT members from innovation schools to also lead particular activities.

Other evidence on how the partnership contributed to achieving district ownership and sustainability comes from the 2016-17 school year, when there was minimal involvement from external partners. Both districts continued supporting PASL/SOAR, and the structures such as network meetings were sustained. A sense of ownership and commitment among individuals was high, even as there were differences between districts in the degree of ownership by district administration. When asked about the extent to which PASL was a priority, a DIDT member in District B replied, "the assistant principal is really supportive. So is our new principal. You know, the whole – of course, the district is –has been very supportive. …The whole social emotional learning initiative is being – at the forefront of our district now." A DIDT member from District A, however, was less positive about the overall district ownership, replying, "District-wide, I would say very little. But at the campuses, you know, especially with campus

administration, they keep that a priority." The next section describes the challenges of moving from individual ownership to districtwide ownership.

From Individual to Districtwide Ownership.

The need to build districtwide ownership was present from the start of the DIDT. Yet in District B, the partnership initially included frequent use of school-based working groups. This appeared to limit the creation of a cohesive districtwide identity. For example, after one school-based activity, a DIDT member in District B asked if the grouping structure meant that they should not be collaborating across groups. In another session, a DIDT member wrote in a feedback form, "there is dissent among the group about our common goal." The program developers (who acted as facilitators in both districts), noted this difficulty and intentionally used more cross-school working groups when the DIDT began its work in District A, and shifted in this direction in District B. There is evidence that this shift was successful, as a DIDT member reflected on the first phase in District A and said, "I think that the presenters have done a good job of making sure that you leave your hat and your title at the door, and everybody is an equal contributing member" (1002).

Despite the increased use of cross-school groups and emphasis on a common district goal, as the SIDTs were created and the work turned more towards focusing on implementation in schools, tensions around whether the work is focused on district or school improvement emerged in both districts. For example, one District A SIDT member reflected on the development phase and said, "Another challenge from last year was just working with — at the point when we had to work with every school, … that was very frustrating, the fact that we really wanted to take what we needed and go our way. Still, being able to give input to everyone else — because I think that's important — but it was frustrating when we had to work together all

three schools on something." This SIDT member is voicing a desire to focus on the important work of implementing the innovation in their school, and aligning it to their specific school's context, and became frustrated with spending too much time collaborating across schools. Similar concerns were expressed in District B as SIDTs were most focused on what their school needed.

As the work continued, there was increasing evidence that SIDTs were more open to learning from each other and greater district cohesion. For example, the scale out phase included multiple instances of school teams engaging with other, with innovation schools sharing their knowledge with scale out schools and all schools discussing common implementation challenges. Near the end of the phase, an innovation school member was asked what was gained from working with other schools; he/she responded, "I've been able to take the meat without the bones. Everything that other groups or schools would share I was able to take it and bring it to my school, and give it to the team to modify and revise based on the needs of our students and the community. They may not have realized it came from this team, but everything we've done on a district level has helped with the work we've done at the school." A principal from a scale out school said he was happy that this initiative had brought the district together: "These meetings, we've become a district…. if you look at these teams here, they aren't [school A, school B], we are [District A], if nothing else, this is what this has done for us. We've got to know each other." His summary was met with loud applause, with an SIDT member saying, "That was amazing, we didn't even need to wrap it up." The importance of this districtwide identity for scaling up was apparent in District B. One DIDT member said,

> […] there is the individuality of each school's needs. But if all three schools are not speaking some sort of commonality, then the goal of scaling out is not going to work. Because B105 needs to say what B106 is saying what B107 is saying […] If we are not saying what it is the same, then we are going to have a problem.

In both districts, even as school teams increasingly worked together around PASL/SOAR, they expressed concern about ownership among district administrators. For school-based DIDT and SIDT members, they were primarily concerned with seeing visible markers of the district ownership in PASL/SOAR. A District A DIDT member said,

> "[My] main concern is that the district doesn't drop the ball. This is something that we want to implement, we will continue district wide, and keep doing instead of just throw it away and try something else…but that's my main concern is that we follow through with this whole process, it's just not dropped, because I want to see it through, and I want to see the change."

Similarly, members in District B wanted to see more district ownership of PASL. At one meeting in the implementation phase, all school coordinators shared that lack of time, and a variety of competing policy initiatives in the district, made implementation of PASL difficult. SIDT member summarized the problem, sharing that "teachers will comply with what they have to comply with, and PASL is not something that they have to comply with […] a lot of different changes are happening right now and PASL is not a priority […]". Several participants reflected that without stronger support from the district, they felt that the idea of scaling PASL within their schools, and to new schools, would likely be difficult. When asked by a researcher what technical assistance the schools need, a DIDT member said,

> From the central office, I think what's key is buy-in from top, so buy-in from the superintendent on down, making sure that he's aware of [this project] and all the work that's being done […] and there's value in it, and then that trickles down to us performing the roles, obviously we feel this is important, that's why we're here but that leadership buy-in is also important.

In District B, many concerns about district commitment was eased midway through the scale out phase, when the superintendent, chief accountability officer, and other senior district leaders attended the DIDT meeting to express commitment to PASL and explain their vision for

PASL in the district. Throughout Phase 4, two influential district leaders took increasing control over the direction of PASL, and stated an intention to eventual spread PASL to "10, 15, and eventually 30 high schools." This was coupled with the creation of a dedicated time for PASL (known as the personalization period). While schools were not mandated to participate in PASL, this district ownership set the conditions for school ownership in District B. At the end of this phase, a DIDT member noted the broad district support for PASL, "From [the superintendent and chief accountability officer] and to the cadre directors have completely bought into the social emotional part of how it's getting students engaged at school. … with the cadre directors… engaged in working with the principals and the core team, you're seeing a greater accountability throughout the schools."

In District A, which was experiencing a prolonged superintendent vacancy and subsequent transition, similar strong statements of long-term commitment in the district were less forthcoming. Further, district administrators wanted to see ownership at the campus level in order for them to support the work, recognizing the challenges that would ensue if they tried to implement SOAR as a mandate. After a researcher says that the scale out phase has largely been a process of the gradual transfer of leadership from the external stakeholders to the district, a central office leader responds, "What [Researcher] said 'as the district takes more role' isn't accurate. If this is going to be maintained, it's going to take active lobbying on the part of the principal to say 'stand with me'. …We need the principals telling us 'come support this.'" Similarly, another central office member praised the ownership he saw from the SIDTs, saying

> Somewhere around year two or three, you no longer needed to hear [district administrators] or anyone else come up here and say, 'You can do it.' In fact, you didn't even ask, you just did it. …You see that switch. You see it. It's not asking for permission anymore. 'I believe I can do this work. I don't want you to tell me what the work is.'

Yet SIDT members were not convinced that this bottom-up ownership for SOAR would be sufficient for scaling up, especially for achieving spread and sustainability. As one SIDT member reflected on the lack of proactive district support and said,

> I don't think this is any kind of malign neglect. I think it's definitely a feeling that here's this program with a bunch of really committed people and what they need is the space to do all the cool things that they're doing, which is probably true to an extent, but there's also an argument to be made that if it's going to be more than just the people who are already in that room, that we need to – we need to actively think about how that next stage could happen.

Another DIDT member agreed, saying, "we have a lot of transition, being a large, urban district, in our schools, that there's not really a good support system that we have of adding new schools or helping the schools that are struggling."

At the school-level, this tension between bottom-up and top-down ownership was also present between principals and teachers. While this was less true in District B, where the SIDT coordinator was also an assistant principal, there was agreement that school administrator commitment was important, with one DIDT member saying, "If you don't have an administrator in that school doing it, it can drift." In District A, however, SOAR became known as an initiative that fostered teacher leadership and ownership. One SIDT member from an innovation school introduced new scale out schools to the work by noting that people from all three innovation schools can "testify very easily to the incredible changes that have been brought to their school through the force of teacher ownership for that change." Despite appreciating feeling empowered, the SIDT members were always concerned with how much their principal supported what they were doing. Two of the innovation schools, for example, mentioned frequently how they struggled to get buy-in from their school administration. A program developer described challenges in one school and said, "teachers need to see the support of their principal, and in that school, the principal didn't really support what the DIDT had come up with… so that –

frustration of kind of having to restart after every session really inhibited, you know, building this ownership and buy in among teachers on the SIDT." Another DIDT member reflected the variation in depth of implementation across schools and noted, "in each of these schools that are really gung-ho and doing really some in-depth good things, they have a principal that's behind them."

Principals, however, described wanting to see ownership among their faculty. For example, one principal in District A, in a meeting to recruit new schools to participate in SOAR, said that it is important it is that principals "give this [work] over to your team." Also, a central office leader noted how often the principals attended the DIDT meetings and said, "I think in the scheme of what principals have to do on a regular basis, the fact that they are here…. that is a huge commitment. …How do we convey to people who don't understand the demands on leadership…that that support IS there?" Overall, then, this type of partnership involves tensions around where strong ownership comes from: teachers, school administration, or district administration.

Establishing Sustainability without Clear Roles or Processes.

While some of these challenges to scaling up in District A were due to the superintendent transition, in both districts there was evidence that the structure of our partnership limited scaling up, particularly for sustainability and depth. One aspect of this comes from the lack of clarity of roles of the DIDT itself and its members after the initial design process. While the DIDT was instrumental in creating the core components of the PASL/SOAR innovation, the development phase saw the DIDT fade into the background as SIDTs took up the work. Both districts saw a decline in attendance and, when they did attend, reduced engagement of at-large DIDT members in during the development phase. For example, fieldnotes from one District A meeting note that

at-large members spent most of the time on their laptops or phones. At-large DIDT members in both districts expressed confusion about what their role is now that the work is focused inside the schools. One DIDT member in District B wrote in a feedback form that while she/he understood the DIDT was to be a "critical friend" of the SIDTs, he/she did not understand what this meant and writing, "I am not quite clear about this. I will assume my input will be valuable and that there will be many opportunities for me to participate with the innovation schools". Another DIDT member said, "the challenges for me as a DIDT member was that at the SIDT team level … I had to really force my way into okay, can you guys keep me updated, as an at-large DIDT team member, I'm not at the school site every day….I wasn't included in a lot of the updates until we came to the convening meeting."

Researcher members of the DIDT were also unsure about their role, with one saying,

> "it was really unclear to me early on how I, as either a researcher or a DIDT member, was supposed to support the SIDT work. … I wasn't really asked to participate in any specific way with regard to that [SIDT kickoff] meeting. I would say that my role was largely to answer any questions around the, that came up by the SIDT members, around the research study from 2010-11, and not that many questions came up."

The lack of clarity of the researcher role on the DIDT was also demonstrated by the unclear role of evidence for DIDT decision-making. While SIDTs used the PDSA process to deepen implementation, when shared with the DIDT members, these presentations were largely about affirming the work of the SIDTs, rather than a critical discussion of evidence. In District B, researchers were pushed by other partners to provide "actionable" information for the school teams. Researchers in both districts provided school-level reports and used this evidence to lead a discussion around implications for future planning during the summer institutes. In network meetings during the year, however, the discussion around this evidence was more informal. For example, a District A DIDT member described what she/he saw as the researcher role: "when

[the research team is] there, they just kind of walk around and listen to what we're talking about and give us some suggestions."

Another way in which the structure of our partnership limited scaling out within the district is the creation of a special team that was solely dedicated to this work. While this helped to foster ownership among individual members as discussed above, there were concerns when the DIDT felt isolated from other parts of the district administration. As a DIDT member in District A noted,

> We used to, at our quarterly meetings, invite in the afternoon district representatives, so our leadership team, our chiefs, if you will, our directors of the schools that the principals answer to; so that they could, you know, not only be aware of what's going on but help out in that process. In that, we really haven't gotten any – any headway in the way we were doing that … They would show up because, you know, they had a calendar evite, you know, and our chief academic officer was asking them to be there. But they never really took ownership, you know, in a participatory type of role.

A notable exception to this lack of engagement by district leaders in District A occurred at the last network meeting, when there was a substantive conversation to develop an "action plan" for sustaining and supporting the scaling up of SOAR. Yet, with few district resources allocated to SOAR, this action plan became a background document used when needed rather than having a visible role in guiding district action. Ultimately, the senior district leaders stopped attending the DIDT meetings altogether, and school-based DIDT members were unsure if this meant they were supposed to take on roles to support scaling up. For example, one member said that while it is nice to hear that senior leadership supports what they are doing,

> it doesn't lend the project the air of something that the district activity would like to promote in other non-SOAR schools. Which leaves the impression that, you know, maybe it is on us to advertise what we're doing to other schools and try to make it attractive to them. But I feel like that's a task that maybe – maybe the district could be helping with more, because it's more of – more of a big picture question.

The DIDT in District B also struggled with its position in the broader district administration. While the DIDT was repeatedly referred to as “critical friends” to the SIDTs and, in later sessions, as “guardians of the design” across the three schools, yet what this meant practically remained unclear. In the implementation phase the lack of integration within the district was apparent, with a researcher saying “we need to build in are we checking in with the district, how does this align with other programs in the district?” The role of the DIDT was described by the developers as shifting to "the central change infrastructure for district-wide scale out". Practically, this shift meant inviting two prominent district leaders to gradually take over leadership of the DIDT and including the innovation school coordinators as key leaders within the network.

Discussion and Conclusion

Looking across the three improvement principles of our approach to scaling up, we see several lessons about achieving ownership, depth, spread, and sustainability, and associated implications for those leading improvement efforts at scale. Focusing on ownership, all three improvement principles helped to build both school and district ownership of the innovations. Involving practitioners in co-constructing the innovation led to a sense of empowerment that they appreciated. Further, the use of continuous improvement cycles evolved to incorporate gathering feedback from teachers, students, and administrators, which also contributed to building ownership among stakeholders. The research-practice partnership structure also fostered local ownership, but it also became clear how ownership across organizational levels is interconnected and difficult to build. These findings are consistent with other scaling up research that emphasizes the need to engage both bottom-up and top-down forms of interaction (Fullan, 2016). Those seeking to scale up effective practices need to develop processes to engage

stakeholders across the organizational level, recognizing that the needs and concerns of school-based personnel will differ from those in the district central office (Cannata et al., in press).

We also identified a dilemma around ensuring adequate depth. While the practitioners involved in this work highly valued the collaborative co-construction process, they also lacked the time and, in some cases, the expertise to design an innovation that would lead to meaningful change in schools. Practitioners in both districts were more likely to support the innovation when they could identify policies, programs or practices already in place at their school and therefore similar to something they were already doing or already had experience doing, or considered a practice relatively easy for teachers to implement. Prior research on achieving scale also emphasizes that reforms are more likely to succeed if they already match the values, goals, and structures of the schools implementing them (Cohen & Mehta, 2017; Hatch & White, 2002). Yet this creates a challenge as deeper engagement with understanding, adapting, and testing the innovation is more likely to build capacity required for improvement (Rubin, Patrick, & Goldring, in press). This dilemma around co-construction, ownership, and depth was also apparent when the teachers in the innovation schools—and not just the SIDT members leading the work in their school—wanted to co-construct the practices for their particular classroom context. Teachers were more removed from the capacity building activities around the innovation than the SIDT members and some of their design changes deviated significantly from the design principles.

The implications for those scaling educational improvement is that this highlights the need for external intermediaries who have both greater expertise around the content of the innovation and more time to focus on development to work in partnership with practitioners. The partner relationship is important as the school and district stakeholders need to trust the external

partners have the expertise and understanding of their local context. External stakeholders can be considered trusted and helpful members of a reform network (Cannata, Redding, Brown, Joshi, & Rutledge, 2017), although this may not always be the case.

In terms of spread, our data identified a dilemma around ownership and spread. Practitioners felt empowered when they had control over the co-construction and continuous improvement testing of the innovation practices in their school. Yet when the work transitioned into on-boarding new schools, this raised questions about what, exactly, was being spread. Spreading the innovation came to be defined as both a set of broad practices that schools adapted to their context, as well as the process of using PDSA to continuously improve those practices in their school. If the district partners only take ownership of the innovation and not the continuous improvement process, effective spread will be inhibited.

Finally, the contextual differences between our districts point to the challenges of building sustainable infrastructure for school improvement. For reforms to last, they need to build a public constituency and district infrastructure to support teachers and schools enacting the innovation (Cohen & Mehta, 2017; Sabelli & Dede, 2013). Infrastructure consists of the organizational structures that provide a forum for the engagement of diverse stakeholders to understand and support each other in the innovation practices (Hopkins & Woulfin, 2015; Scherrer, Israel, & Resnick, 2013). Yet our findings indicate that building these structures requires ownership and commitment from senior district leaders, which can be disrupted with substantial leadership turnover (Finnigan, Daly, & Liou, 2016).

Our findings also highlight how the dimensions of scale are not linear. Rather, they ebb and flow over time. A sense of ownership can be high, then low, and then high again. The

dimensions evolve over time, as the innovation that is being scaled itself can evolve (Dede, 2006).

There is a recent focus on research-practice partnerships that more effectively address the needs of practitioners when implementing new initiatives to achieve implementation integrity, depth, sustainability, and scale (Coburn & Penuel, 2016). The emergence of design-based implementation research reflects this need for researchers to work in partnership with practitioners to design, implement, and scale educational innovations (Penuel et al., 2011). By both enacting a new approach to scale and studying how the key principles of the approach contribute to scale, this paper provides crucial knowledge about how improvement science and design-based implementation research may lead to meaningful change in schools.

References

Bradach, J. (2003). *Going to Scale: The Challenge of Replicating Social Programs* (Stanford Social Innovation Review). Stanford, CA: Stanford University. Retrieved from https://ssir.org/images/articles/2003SP_feature_bradach.pdf

Bryk, A. S., Gomez, L., Grunow, A., & Hallinan, M. T. (2011). Getting Ideas into Action: Building Networked Improvement Communities in Education. In *Frontiers in Sociology of Education*. Springer Publishing.

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 Education Press.

Bryk, A. S., Yeager, D. S., Hausman, H., Muhich, J., Dolle, J. R., Grunow, A., … Gomez, L. (2013). *Improvement Research Carried Out Through Networked Communities: Accelerating Learning about Practices that Support More Productive Student Mindsets* (pp. 1–32). Carnegie Foundation for the Advancement of Teaching.

Cannata, M., Cohen-Vogel, L., & Sorum, M. (in press). Partnering for Improvement: Improvement Communities and Their Role in Scale-Up. *Peabody Journal of Education*.

Cannata, M., Redding, C., Brown, S., Joshi, E., & Rutledge, S. A. (2017). How Ideas Spread: Establishing a Networked Improvement Community. In *Presentation at the Carnegie Foundation Summit on Improvement in Education*. San Francisco, CA.

Cannata, M., & Rutledge, S. A. (in press). New Frontiers in Scaling Up Research. *Peabody Journal of Education*.

Chambers, D. A., Glasgow, R. E., & Stange, K. C. (2013). The dynamic sustainability framework: Addressing the paradox of sustainment amid ongoing change. *Implementation Science*, *8*, 117.

Coburn, C. E. (2003). Rethinking Scale: Moving Beyond Numbers to Deep and Lasting Change. *Educational Researcher*, *32*(6), 3–12. https://doi.org/10.3102/0013189X032006003

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., Penuel, W. R., & Geil, K. E. (2013). *Research-Practice Partnerships: A Strategy for Leveraging Research for Educational Improvement in School Districts*. New York, NY: William T. Grant Foundation.

Coburn, C. E., & Turner, E. O. (2011). Research on Data Use: A Framework and Analysis. *Measurement: Interdisciplinary Research and Perspectives*, *9*(4), 173–206. https://doi.org/10.1080/15366367.2011.626729

Cohen, D. K., & Mehta, J. D. (2017). Why Reform Sometimes Succeeds: Understanding the Conditions That Produce Reforms That Last. *American Educational Research Journal*, 0002831217700078. https://doi.org/10.3102/0002831217700078

Cohen, D. K., Peurach, D. J., Glazer, J. L., Gates, K. E., & Goldin, S. (2013). *Improvement by Design: The Promise of Better Schools*. Chicago ; London: University Of Chicago Press.

Cohen-Vogel, L., Cannata, M., Rutledge, S., & Socol, A. R. (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, L., Tichnor-Wagner, A., Allen, D., Harrison, C., Kainz, K., Socol, A. R., & Wang, Q. (2015). Implementing Educational Innovations at Scale Transforming Researchers Into Continuous Improvement Scientists. *Educational Policy*, 0895904814560886.

Corbin, J. M., & Strauss, A. (2008). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory.* (3rd ed.). Los Angeles, CA: Sage Publications, Inc.

Cuban, L. (1998). How Schools Change Reforms: Redefining Reform Success and Failure. *Teachers College Record*, *99*(3), 453–77.

Datnow, A. (2005). The Sustainability of Comprehensive School Reform Models in Changing District and State Contexts. *Educational Administration Quarterly*, *41*(1), 121–153. https://doi.org/10.1177/0013161X04269578

Datnow, A., & Park, V. (2009). Conceptualizing policy implementation: Large-scale reform in an era of complexity. In G. Sykes, B. Schneider, & D. N. Plank (Eds.), *Handbook of Education Policy Research* (1 edition, pp. 348–361). New York : Washington, D.C.: Routledge.

Dede, C. (2006). Scaling up: Evolving innovations beyond ideal settings to challenging contexts of practice. In R. K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 551–566). Cambridge, UK: Cambridge University Press.

Desimone, L. (2002). How Can Comprehensive School Reform Models Be Successfully Implemented? *Review of Educational Research*, *72*(3), 433–479. https://doi.org/10.3102/00346543072003433

Durlak, J. A., & Dupre, E. P. (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*, 327–350.

Elmore, R. F. (2016). “Getting to scale…” it seemed like a good idea at the time. *Journal of Educational Change*, *17*(4), 529–537. https://doi.org/10.1007/s10833-016-9290-8

Finnigan, K. S., Daly, A. J., & Liou, Y.-H. (2016). How leadership churn undermines learning and improvement in low-performing school districts. In A. J. Daly & K. S. Finnigan (Eds.), *Thinking and acting systemically: Improving school districts under pressure* (pp. 183–205). Washington, D.C.: American Educational Research Association.

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

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: theories, methods, and exemplars* (pp. 136–156). New York: Teachers College, Columbia University.

Fullan, M. (2016). The elusive nature of whole system improvement in education. *Journal of Educational Change*, *17*(4), 539–544. https://doi.org/10.1007/s10833-016-9289-1

Furr, N. R., & Ahlstrom, P. (2011). *Nail It then Scale It: The Entrepreneur’s Guide to Creating and Managing Breakthrough Innovation* (First Edition, June 2011 edition). United States? NISI Institute.

Glennan, T. K., Bodilly, S. J., Galegher, J. R., & Kerr, K. A. (2004). *Expanding the Reach of Education Reforms: Perspectives from Leaders in the Scale-Up of Educational Interventions* (1st ed.). Rand Publishing.

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

Hatch, T., & White, N. (2002). The Raw Materials of Reform: Rethinking the Knowledge of School Improvement. *Journal of Educational Change*, *3*(2), 117–134. https://doi.org/10.1023/A:1016516212204

Hopkins, M., & Woulfin, S. L. (2015). School system (re)design: Developing educational infrastructures to support school leadership and teaching practice. *Journal of Educational Change*, *16*(4), 371–377. https://doi.org/10.1007/s10833-015-9260-6

Langley, G. J. (2009). *The improvement guide: a practical approach to enhancing organizational performance*. San Francisco: Jossey-Bass.

LeMahieu, P. G., Grunow, A., Baker, L., Nordstrum, L. E., & Gomez, L. M. (2017). Networked Improvement Communities: the discipline of improvement science meets the power of networks. *Quality Assurance in Education*. https://doi.org/10.1108/QAE-12-2016-0084

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

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

McLaughlin, M. W. (1976). Implementation as Mutual Adaptation: Change in Classroom Organization. *Teachers College Record*. Retrieved from http://www.eric.ed.gov/ERICWebPortal/detail?accno=EJ135285

Means, B., & Penuel, W. R. (2005). Scaling Up Technology-Based Educational Innovations. In C. Dede, J. P. Honan, & L. C. Peters (Eds.), *Scaling Up Success : Lessons Learned from Technology-Based Educational Improvement* (1st ed., pp. 176–197). Jossey-Bass.

Penuel, W. R., Fishman, B. J., Cheng, B. H., & Sabelli, N. (2011). Organizing Research and Development at the Intersection of Learning, Implementation, and Design. *Educational Researcher*, *40*(7), 331–337. https://doi.org/10.3102/0013189X11421826

Peurach, D. J. (2011). *Seeing Complexity in Public Education: Problems, Possibilities, and Success for All* (1st ed.). Oxford University Press, USA.

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.

Peurach, D. J., & Neumerski, C. M. (2015). Mixing metaphors: Building infrastructure for large scale school turnaround. *Journal of Educational Change*, *16*(4), 379–420. https://doi.org/10.1007/s10833-015-9259-z

Redding, C., Cannata, M., & Taylor Haynes, K. (in press). With Scale in Mind: NCSU's Integrated Model of School-Based Design and Implementation. *Peabody Journal of Education*.

Rubin, M., Patrick, S., & Goldring, E. (in press). Make it quick or make it last? Dilemmas in the implementation of a school improvement reform. *Peabody Journal of Education*.

Rutledge, S. A., & Cannata, M. (2016). Identifying and Understanding Effective High Schools. *Phi Delta Kappan*, *97*(6), 60–64.

Sabelli, N., & Dede, C. (2013). Empowering design-based implementation research: The need for infrastructure. In B. J. Fishman, W. R. Penuel, A.-R. Allen, & B. H. Cheng (Eds.), *Design-based implementation research: theories, methods, and exemplars* (pp. 464–480). New York: Teachers College, Columbia University.

Sabelli, N., & Harris, C. J. (2015). The Role of Innovation in Scaling Up Educational Innovations. In C.-K. Looi & L. W. Teh (Eds.), *Scaling Educational Innovations*. Singapore: Springer.

Scherrer, J., Israel, N., & Resnick, L. B. (2013). Beyond classrooms: Scaling and sustaining instructional innovations. In B. J. Fishman, W. R. Penuel, A.-R. Allen, & B. H. Cheng (Eds.), *Design-based implementation research: theories, methods, and exemplars* (pp. 426–442). New York: Teachers College, Columbia University.

Siskin, L. S. (2016). Mutual Adaptation in Action. *Teachers College Record*, *118*(13), 1–18.

Sowers, N., & Yamada, H. (2015). *Pathways Impact Report*. Stanford, CA: Carnegie Foundation for the Advancement of Teaching. Retrieved from http://cdn.carnegiefoundation.org/wp-content/uploads/2015/01/pathways_impact_report_2015.pdf

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

Sutton, R. I., & Rao, H. (2014). *Scaling Up Excellence: Getting to More Without Settling for Less* (1 edition). New York: Crown Business.

Thompson, M., & Wiliam, D. (2008). Tight but loose: A conceptual framework for scaling up reforms. In E. C. Wylie (Ed.), *Tight but loose: Scaling up teacher professional development in diverse contexts* (pp. 1–44). Princeton, NJ: ETS.

Tables and Figures

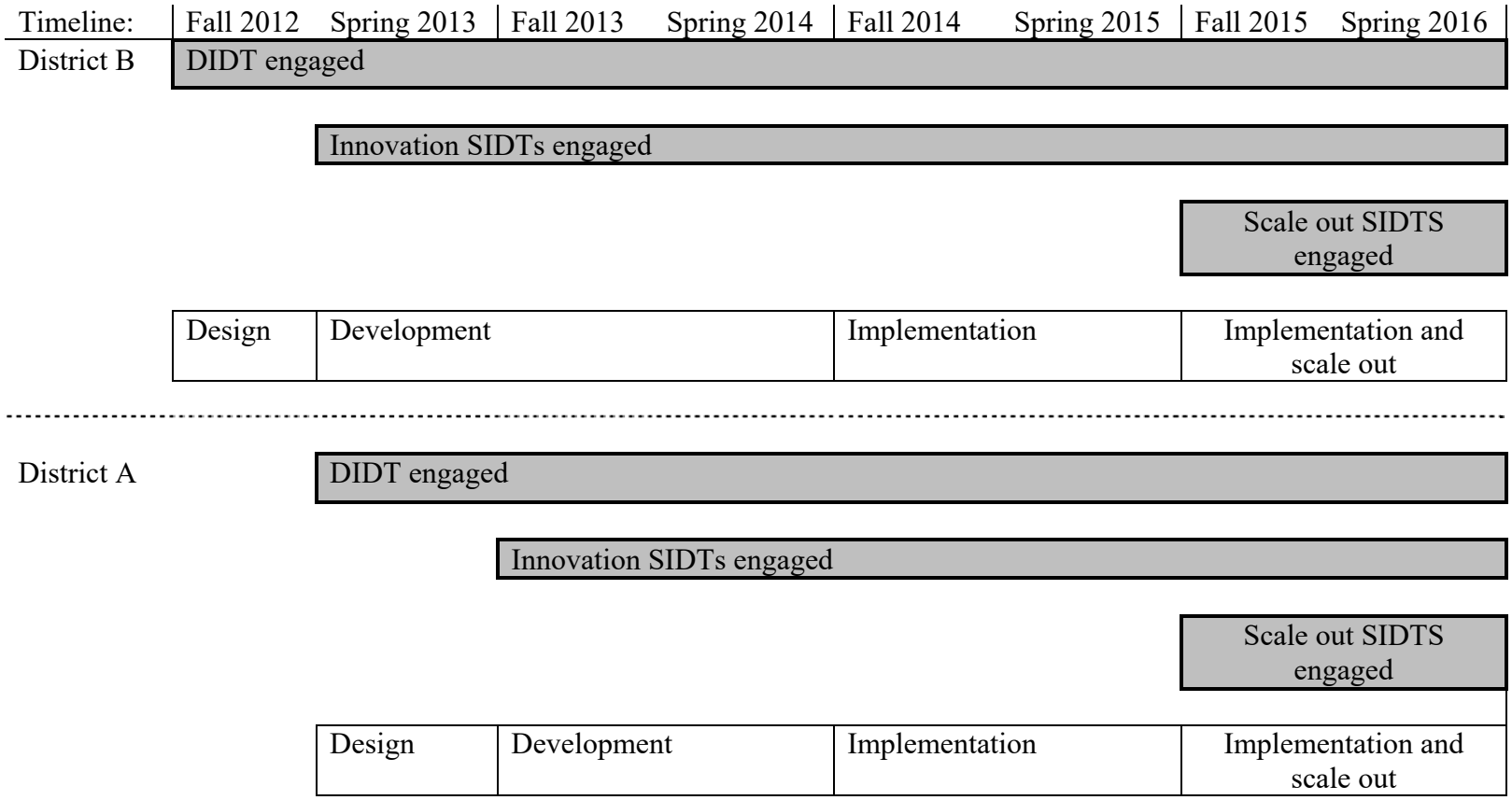

Figure 1 - Timeline of partnership activities

Table 1 –Interview and Focus Group Data Collected In Each District

	Spring 2013	Summer 2013	Fall 2013	Summer 2014	Fall 2014	Spring 2015	Fall 2015	Spring 2016	Fall 2016
District A									
DIDT/SIDT interviews		21	24	16	20	22	14	57	10
Administrator interviews			9		11	12		9	
Teacher interviews			5		71	70		67	
Student focus groups			9		12	14		20	
Teacher focus groups			8		8	5		0	
Other personnel interviews			10		0	1		0	
Other personnel focus groups			0		3	3		0	
District B									
DIDT/SIDT interviews	23	16	3	20		18	15	30	14
Administrator interviews	11					6		8	
Teacher interviews	48					35		58	
Student focus groups	9					9		15	
Teacher focus groups	0					0		7	
Other personnel interviews	8					0		11	