Continuous Improvement in the Public School Context:

Understanding How Educators Respond to Plan-Do-Study-Act Cycles

Ariel Tichnor-Wagner
John Wachen
Marisa Cannata
Lora Cohen-Vogel

April 2017

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

This is a post-peer-review, pre-copyedit version of an article published in the Journal of Educational Change. The final authenticated version is available online at: https://doi.org/10.1007/s10833-017-9301-4

Abstract

The last five years have witnessed growing support amongst government institutions and educational foundations for applying continuous improvement research (CIR) in school settings. CIR responds to the challenge of implementing effective educational innovations at scale by working with practitioners in local contexts to understand "what works, for whom, and under what conditions." CIR works to achieve system improvement through the use of Plan-Do-Study-Act (PDSA) cycles, which are multiple tests of small changes. This comparative case study of two urban school districts examined how innovation design teams took up PDSA in their work to improve high school student outcomes, and their perceptions of PDSA as an approach to innovation development, adaptation, and implementation. Findings revealed both possibilities and challenges for implementing PDSA. Nearly all participants reported the value in PDSA, and participants pointed to *connections to previous experiences* and *PDSA training* as helping to build capacity. However, we found mixed levels of enthusiasm for actually conducting PDSA cycles, and capacity constraints regarding time and data collection.

Continuous Improvement in the Public School Context:
Understanding How Educators Respond to Plan-Do-Study-Act Cycles

The last five years have witnessed growing support amongst government institutions and educational foundations for applying continuous improvement research (CIR) in school settings. CIR responds to the challenge of implementing effective educational innovations at scale by working with practitioners in local contexts to understand "what works, for whom, and under what conditions" (Identifying Reference; Park et al., 2013). Proponents of CIR argue that real, lasting improvement requires combining scientific discipline and practitioner knowledge of local context across a diverse set of school conditions (Bryk, 2009; Lewis, 2015).

Theoretically and in practice, CIR works to achieve system improvement through the use of multiple tests of small changes (Morris & Hiebert, 2011). These tests rely on a Plan-Do-Study-Act (PDSA) cycle, which guides practitioners to set measurable aims and test whether the changes they make result in improvement (Deming, 2000). During PDSA cycles, practitioners "Plan" a change to be tested, "Do" the test, "Study" the data they collected during the test, and "Act" on what they have learned from the test by abandoning, revising, or scaling up the change (Langley et al., 2009).

Fields including industry and healthcare began using continuous improvement approaches in earnest more than two decades ago, and have, since that time, demonstrated improved outcomes as a result (e.g., Bheuyan & Baghel, 2005; Deming, 2000). Increasingly, the PDSA framework is being used in educational settings; early indications suggest that in education, too, CIR is showing promise. For example, in the Middle School Mathematics and the Institutional Setting of Teaching (MIST) project, researchers worked with practitioners in four districts to establish an empirically grounded theory of action for improving the quality of mathematics instruction at scale (Cobb et al., 2013). The improvement approach led to new

decision-making routines in the districts and to robust instructional improvements in mathematics among teachers. The Carnegie Foundation for the Advancement of Teaching's Pathways program, an initiative that uses a CIR approach to solve the problem of low developmental math completion rates, has also shown promising results. Through Pathways, networks of educators across 19 community colleges in five states designed and tested a new approach for teaching developmental mathematics. In the first three years of implementing two new courses, Statway and Quantway, approximately 50% of students enrolled in the courses successfully completed the pathways, which is substantially higher than the typical completion rates of students in traditional developmental math sequences (Sowers & Yamada, 2015).

Despite these early successes, little is known about the viability of widespread use of CIR in K-12 public school settings. Drawing on data from our continuous improvement work with two large, urban school districts as part of [identifying study], we seek to understand how school and district personnel, researchers, and development specialists respond to the use of PDSA cycles to design and test educational innovations that emphasized both rigorous academic and social-emotional learning for high school students. We examine how, when working in innovation design teams to implement improvement cycles for the first time, practitioners and development partners perceived their will and capacity to carry out PDSA cycles. Specifically, we ask: 1) What was the will of innovation design team members to implement PDSA, and what aspects of will impacted how they engaged? 2) What was the capacity of innovation design team members to implement PDSA, and what aspects of capacity impacted how capacity was built? 3) How do these aspects of will and capacity compare across the two districts?

Before answering these questions, we begin with an overview of the CIR approach and its origins. The following section describes the key feature of CIR: the Plan-Do-Study-Act cycles

used to test changes in real world settings. Next, we discuss *will* and *capacity,* key tenets from the implementation literature, and why we believe their study is important in determining to what extent educators are likely to adopt CIR to solve problems of practice. In the methods section, we describe the research settings and data we collected in order to examine participants' will and capacity to use PDSA. We include in that section a description of the learning opportunities (i.e., training and resources) participants received on PDSA. We then present the case of PDSA implementation in each district, followed by a cross-case comparison that draws out the commonalities and contrasts in PDSA implementation between the two districts.

The 'Science of Improvement': What Is It and From Where Did It Come?

In a 2011 *Education Week* article, Anthony Bryk of the Carnegie Foundation proposed a "science of improvement" – a productive synthesis, in his view, of the strengths of randomized control trials and action research. The improvement approach is distinct from traditional forms of quantitative research in at least four ways. First, whereas conventional research in the quantitative tradition often aims to hold variables constant and uses a set of fixed procedures to carry out the work, CIR "focuses on characterizing the situation in all its complexity" and uses an iterative, flexible process wherein design and research plans are revised as the work progresses (Barab & Squire, 2004, p. 4; Lewis, 2015). Second, whereas traditional research primarily focuses on outcomes of interventions, CIR also involves study of the design process; there is an interest, that is to say, in understanding the workings of the approach itself in an effort to improve it the next time (Bryk et al., 2015). Third, whereas traditional research "uses well-validated tools to measure implementation and impact," CIR utilizes "practical measurement tools" that test the indicators predicted in the theory of change (Lewis, 2015, p. 56). Such measures provide quick feedback into the system, using a small number of targeted questions.

Fourth is a reconceptualization of the researcher's role. Although pains are often taken in traditional quantitative research to keep researchers "outside" the intervention being tested, CIR purposely involves the researchers in innovation design and revision. Moreover, researchers "are expected to become smarter about how to target issues that matter" to educators and about "how to conduct solid research within the constraints of practicing education systems" (Means & Harris, 2013, p. 360). The role of participants in the research is different too, aligning more with the tradition of action research: "Whereas traditional research often treats participants (e.g., teachers) as subjects, improvement science includes participants in the design process, involving them as equals in the work" (Identifying Reference, 2015, p. 12).

While CIR is new to education, it has roots in business and healthcare (Bheuyan & Baghel, 2005). Attributed to the 20th century statistician W. Edwards Deming, the improvement approach was further developed in healthcare by Donald Berwick (Mann, 1993), who co-founded the Institute for Healthcare Improvement (IHI) in 1989. Since its inception, IHI has driven health care improvement research; today, its website contains links to more than 1,000 publications related to improvement in health care delivery and outcomes (Shortell et al., 1998). To help explain a surge in the use of the improvement approach in public health and industry, Berwick points to Associates in Process Improvement (API) and its *Improvement Guide: A Practical Approach to Enhancing Organizational Performance* (see Langley et al., 2009). The publication of *The Improvement Guide* (Langley et al., 2009), a comprehensive description of the frameworks, tools, and processes of improvement science, has also fueled the use of CIR among educators and educational researchers (Identifying Reference, 2015; Lewis, 2015).

Another catalyst for the adoption of improvement models in education is the use of designed-based implementation research, a close cousin of CIR, which focuses on design and

testing in local contexts and emphasizes collaboration between researchers and practitioners (Anderson & Shattuck, 2012; Fishman et al., 2013). Further, CIR and design-based research build on the tradition of action research, wherein practitioners use data collection and analysis to evaluate how to improve their practice. Action research contains "a spiral of self-reflective cycles" that include the following: planning a change, acting and observing the change process and results, reflecting on the process and results, and re-planning (Kemmis & McTaggart, 2005).

CIR in practice can vary substantially, depending on discipline, context, goals, and other characteristics of the research project. For example, "Networked Improvement Communities" (NICs) are at the core of the approach to CIR utilized by the Carnegie Foundation for the Advancement of Teaching. NICs are professional networks that facilitate the transfer of information across contexts, thereby allowing for greater collective knowledge (Dolle, Gomez, Russell, & Bryk, 2013). Other approaches emphasize design-based implementation research, where teams focus on identified problems of practice and use iterative, collaborative design to understand both the problem and potential solutions, but also how to implement these solutions (Fishman, Penuel, Allen, & Cheng, 2013). Most enactments of CIR take the form of research-practitioner partnerships, in which the focus is on problems that practitioners, in collaboration with researchers, identify as relevant to practice. Rather than focusing on gaps in research or theory, research priorities explicitly respond to practitioner needs (Penuel, Allen, Coburn, & Farrell, 2015; Tseng, 2012).

The approach utilized in the partnership described in this study has elements of these other approaches. Although a researcher-practitioner partnership was a central aspect of the work, the initial research phase of the work did not directly involve practitioners in identifying problems of practice and research priorities. Instead, researchers collected and analyzed data in

the participating districts and identified effective practices across higher-performing schools that were then shared with the practitioner teams. The collaborative work began with the development of specific innovations based on these effective practices. The "Setting and Context of PDSA" section below provides additional information about the characteristics of the project reported on in this study.

Plan-Do-Study-Act Cycle: The Crux of Continuous Improvement

At the crux of CIR is the Plan-Do-Study-Act (PDSA) cycle. PDSA cycles test a change in real world settings and help improvement teams determine whether a change is an improvement (Langley et al., 2009). Similar to action research, a PDSA cycle consists of four parts. First, the improvement team *plans* the test, asking what change ("prototype") will be tested, with whom/with what measures it will be tested, and what changes are expected as the result of trying out the prototype. Second, the team *does* the test, gathering information on what happened during the test and as a result of it. Third, the team *studies* the information gathered during the test, comparing it with predictions made about the prototype's effects. Having studied the information, the team *acts*, making a decision about whether to abandon the prototype, revise it, or scale it up with a larger number of users.

After testing the change on a small scale—with a few teachers or classrooms—PDSA cycles repeat (see Figure 1). The improvement team learns from each test, refines the change, and then may implement the change on a broader scale—for example, with an entire grade level. After successful implementation within a unit, the team can continue to use PDSA to spread, or bring to scale, the change to other parts of the organization or other organizations entirely, effectively using it to adapt the change to new contexts, resulting in system change. This kind of

cycled inquiry can take various forms (e.g., RCTs, quasi-experimental designs) and affords adaptation to context or what some refer to as "local tests of change" (Park et al., 2013).

Lengths of test cycles vary. Although short (90-day) cycles are common, improvement science allows for short- and long-cycle testing, depending on the research questions and nature of the change itself. By encouraging early and iterative testing of ideas in the specific environment of interest, the improvement model allows the innovation to be gradually modified to the uniqueness of the system in which it is being implemented (Langley et al., 2009).

In an example from [identifying reference], *innovation design teams* in Broward County, Florida used findings from a comprehensive study of their high schools to design an innovation prototype they called PASL, or Personalization for Academic and Social-Emotional Learning. Among its other components, the PASL innovation would assign every 9th grade student to a PASL teacher who was responsible for conducting "rapid check-ins" (RCIs) with these students. Rapid check-ins consisted of very brief, impromptu interactions between a PASL teacher and her assigned students, during which the teacher would simply 'check in' with the student about her day, how her sports team fared last night, or whether she is looking forward to Homecoming, for example. Design team members expected RCIs to build teachers' knowledge about the interests and concerns of their students and, as a result, improve students' sense of connection to the school.

To test RCIs, during the "Plan" phase, the innovation design teams at three study schools identified a small subset of teachers with whom to try them out. The teams assigned these teachers to students in one of their class periods and gave teachers an RCI log to record the number of rapid check-ins they conducted with each student over a two-week period. After the two weeks of testing during the "Do" phase ended, innovation design team members met with

the implementing teachers to "Study" the RCI logs and gather feedback from teachers. Feedback indicated the logs were too cumbersome, and the innovation team "Acted" to simplify the form and try out the RCIs with a larger number of teachers.

In this study, the improvement teams were introduced to the PDSA approach with two objectives in mind: 1) to further develop, refine, implement, and scale the designed components of an innovation, and 2) to build the capacity of schools and the district to engage in continuous improvement for future innovations they might implement in the future. Rather than examining whether PDSA led to improved teacher and student outcomes, we analyzed how practitioners and development partners working in high schools respond to the challenge of utilizing the PDSA approach by examining their perceived will and capacity to engage in this type of improvement process and comparing these aspects of will and capacity across districts.

Will and Capacity: A Framework for Examining Perceptions of PDSA

As a relatively new approach in education, it is important to understand educators' perceptions of implementing PDSA. This is particularly important as the evidence from healthcare has pointed to challenges associated with implementing the four components of PDSA in a cyclical fashion (Ogrinc & Shojania, 2014; Taylor et al., 2014). One challenge has been attitudinal, for example, healthcare professionals not seeing the value of documentation during the "Do" and "Study" phase (Reed, 2015) and having issues adjusting their perception of rigor to encompass more than RCTs (Berwick, 1998). Another set of challenges has related to healthcare professionals' capacity to implement PDSA cycles, including time and money constraints (Berwick, 1998) and the frequency, relevance, and reliability of data collected to inform the "Study" and "Act" phases (Reed, 2015; Tyler et al., 2014).

These attitudinal and ability-related challenges reflect what implementation research in education has long noted, "what matters most to policy outcomes are local capacity and will" (McLaughlin, 1990, p. 12). *Will* is the motivation to embrace reform objectives, which includes believing in the value of a reform (e.g., believing that a new curricular framework will lead to student improvements) and an eagerness "to commit energy and effort to its success" (McLaughlin, 1990, p. 13). *Capacity* is the knowledge, skills, organizational routines, resources, and personnel available to support implementation (Firestone, 1989; McLaughlin, 1990; Spillane et al., 2002). An individual's capacity to enact a given reform may be influenced by their prior knowledge and experiences, which may cloud one's understanding of what a reform is asking them to do. Implementers may interpret the same message in different ways, misinterpret new ideas as familiar, or focus on superficial features rather than deeper differences between current knowledge or practices and new reforms (Spillane et al., 2002). This may lead to a lack of ability to implement a reform as the design intended, even when implementers express the will to do so.

Traditional implementation research has largely portrayed a lack of local will and capacity as impediments to reform (Firestone, 1989; Spillane et al., 2002). Yet will and capacity are not inherent, fixed conditions; both can be shaped by the learning opportunities that implementers experience (Cohen & Hill, 2001) and the organizational and political contexts in which they work (McLaughlin, 1990). Therefore, we first explain the district context and the PDSA learning opportunities that practitioners experienced before turning to our findings on the will and capacity of high school practitioners to implement PDSA. Given that will and capacity influence the extent to which new processes and programs are implemented, which in turn mediates the efficacy of those processes and programs, these findings serve as a vital first step in

ascertaining whether using PDSA cycles in school contexts lead to replicable, scalable innovations.

Methods

Drawing upon data from a multi-year research project conducted by [identifying reference], we employ a comparative case study (Yin, 2009) of innovation design teams in two large, urban districts engaging in improvement work. Members of the innovation design teams included district central office and high school practitioners, researchers, and development specialists. During the 2012-2013 school year, the innovation design team participated in a "design challenge," which charged them with developing an innovation based on the core elements of a set of practices that researchers identified in year-long case studies of four high schools in each of two respective districts (See Identifying Reference for more about the schools themselves). This innovation would first be implemented in three innovation high schools and eventually scaled up to other high schools across the district. As such, the practitioners involved in the innovation design team for each district included teacher and administrative representatives from the three innovation schools, at-large members from other high schools in the district, and central office administrators. The two innovation design teams began implementing PDSA in the 2013-2014 school year to "test" components of the innovation prototype in the three innovation high schools. This study examined the design team members' perceptions of PDSA during this first year of PDSA implementation.

Setting and Context of PDSA

District context. We hypothesized that, similar to innovations or policies, the implementation of the PDSA approach would be influenced by the contexts of the two districts. Specifically, educators' responses to and perceptions of the approach may be influenced by a

variety of contextual factors, including previous experiences with CIR (or CIR-like approaches), accountability pressures, and other district-level structures and policies. Table 1 provides an overview of the Broward and Fort Worth district contexts where PDSA implementation occurred. Both districts are large, culturally-, linguistically-, and economically-diverse urban school districts in states with long histories of high stakes accountability. Both districts had also recently incorporated continuous improvement into their strategic plans. For example, Fort Worth adopted the Baldrige Excellence Framework, which involves a systems approach to organizational improvement and includes a PDSA model, shortly before PDSA was introduced as part of this work. However, the use of PDSA cycles as a part of an improvement process was novice to the practitioners on both districts' innovation design teams.

The improvement work was organized similarly in the two districts, with the innovation design team responsible for designing and developing the innovation and, subsequently, adapting the innovation to each school context through PDSA cycles. The innovation design teams were similarly composed across the districts, although the Fort Worth teams had slightly more representation from classroom teachers. In addition to the school and district practitioners, researchers, and development specialists, each district had a liaison who helped coordinate activities in the district and served as a bridge between the researchers/development specialists and practitioners.

While following similar improvement processes, the two districts designed innovations that were unique and responsive to findings from case studies conducted by researchers in each district. In Broward, the innovation focused on Personalization for Academic and Social Emotional Learning (PASL). In Fort Worth, the innovation focused on building Student Ownership And Responsibility (SOAR). Table 1 enumerates the components of each innovation.

PDSA learning opportunities. As illustrated in Figure 2, the Broward and Fort Worth design teams received similar PDSA trainings during the 2013-2014 school year. Researchers and development specialists led PDSA trainings for school and district practitioners, with researchers delivering most of the content and development specialists providing most of the technical support. Prior to leading these PDSA trainings, the researchers and development specialists received training in continuous improvement and PDSA cycles through the Carnegie Foundation for the Advancement of Teaching. Yet this was the first time they led PDSA training.

These formal PDSA learning opportunities occurred about once a month, taking the form of face-to-face sessions wherein researchers, development specialists, and practitioners met in person for one to two full days; two-hour cross-district webinars and within-district webinars that took place after school; check-and-connect calls between school coordinators overseeing PDSA, researchers, and development specialists; and a summer institute. Learning opportunities did not focus exclusively on PDSA; during a typical innovation design team session, time was split between discussing PDSA, the innovation design, and planning for implementation. In addition, the researchers provided the district and school practitioners with a PDSA facilitation guide in the fall of 2013. This 32-page document served as a reference tool that defined and provided examples of each PDSA stage and provided forms for implementers to fill out as they completed each phase of the cycle. For illustration purposes, one such form, adapted from Langley et al. (2009), is included in Appendix A.

The learning opportunities in both districts followed a similar timeline. In the fall of 2013, researchers introduced each stage of the PDSA cycle and how it would be a part of the ongoing innovation development and implementation processes. Researchers guided school practitioners through the first testing cycle in late fall and early winter. For example, after

researchers introduced the PDSA cycle, they worked with school practitioners as they planned what their first cycle would entail, helping them understand what was needed in the "Plan" phase and the tools that were available in the PDSA facilitation guide to support the "Study" phase. In Broward, all innovation schools tested a professional development module on Rapid Check-Ins (RCIs). In Fort Worth, each innovation school chose a different change idea to test; one school tested a behavioral reflection sheet, one school tested teacher professional development on growth mindsets and problem-solving, and one school tested using praise language that fosters a growth mindset.

In meetings during the spring of 2014, the design teams engaged in the "Plan", "Study", and "Act" phases of the PDSA cycle all in one sitting. For example, during the February webinar in Broward the team partook in the "Study" and "Act" phase of the first cycle's test of the professional development on RCIs and planned for the second cycle. This cycle, rather than build on the previous cycle testing RCI professional development, tested a different aspect of the PASL innovation: the RCI collection form with a handful of PASL teachers. During the March face-to-face meeting, the innovation design team studied the PDSA Cycle 2 on RCIs and planned Cycle 3. During the spring meetings, researchers in both districts led an in-depth discussion of collecting, analyzing, and using data within PDSA cycles. In Fort Worth, the March face-to-face session included time for the teams to "Study" and "Act" on their first cycle, share results across schools, and time to "Plan" their second and third cycles. As school teams shared the results from their first cycle, two of the schools were impressed with the results from the third school's test of a behavioral reflection sheet. These schools "Planned" their second cycle to test a similar practice. The "Plan" for the third school focused on recruiting additional teachers to test the

practices they developed. In the May session, the teams had time to "Study" and "Act" on their second cycle.

Data Sources

Data sources included interviews with members of the innovation design teams who participated in PDSA training and implementation, along with documents and field notes from those trainings. We used semi-structured interviews as our primary data source. Interviews were conducted with district and school practitioners responsible for developing and designing the innovation (n = 35), researchers (n = 6), and development specialists (n = 5). These interviews were conducted during the summer of 2014, after the first year of PDSA implementation. During interviews, participants were asked about their role in undertaking PDSA, the value of PDSA, what they would have done differently if conducting a PDSA cycle again, and their capacity to continue to conduct PDSA cycles.

In addition, we drew upon observations and artifacts from PDSA training sessions (n = 19) across both districts between the summer of 2013 and the summer of 2014. These data included field notes taken by the researchers during 12 innovation design team meetings in the districts, 14 researcher reflection forms, handouts and PowerPoint presentations delivered to practitioners about PDSA, and feedback forms completed by participants following four of these meetings. Participant feedback forms included both Likert-type scales and open-response items, asking, for example, "Do you feel that the process of PDSA testing helped your team to improve Rapid Check-Ins?" and "Please describe what lessons your team learned from the PDSA cycles."

Data Analysis

Data analysis consisted of an initial round of descriptive coding that captured examples of the *local context* and *learning opportunities* surrounding PDSA and participants' *will* and

capacity for implementing it. From this first round of coding, a subset of codes emerged from the data that represented themes under the broad constructs of will and capacity (Miles & Huberman, 1994). The researchers conducted a second round of coding of all data sources to include the sub-codes that emerged. Sub-codes under will included *enthusiasm*, *compliance*, *resistance*, *value*, and *paperwork.* Sub-codes under capacity included *connections to similar practices, data collection/analysis, documentation, trainer expertise, training, previous PDSA experience,* and *time/resources.*

To ensure reliability, the coding process began with two researchers coding the data. First, both researchers coded the same set of interviews and met on a weekly basis to compare codes and address disparate coding examples and areas of confusion. Incongruence was addressed by modifying code definitions and including examples of each code in the codebook. Inter-rater reliability was established when both researchers coded the same text with 90% reliability, after which researchers coded the remaining data. The coding team continued to meet on a weekly basis to discuss emerging codes, themes, and patterns.

Using the coded data, in-depth analytic memos were written for each district. A cross-district matrix was then created to compare findings between districts, drawing out similarities and differences in how participants discussed will and capacity to engage in PDSA (Miles & Huberman, 1994). Throughout this process, we searched for areas of convergence and divergence between sub-codes (Marshall & Rossman, 2011), merging overlapping codes and ensuring individual codes were distinct from one another. For example, we merged *trainer expertise* and *training* together, and did the same for *resistance* and *paperwork*. This eventually led to the themes that illuminated the will and capacity of innovation design team members to implement PDSA, discussed in the following section.

Findings: Perceptions of Implementing PDSA in High Schools

Overall, similar themes emerged across both districts that explained high school practitioners' will and capacity to implement PDSA during the first year of innovation development. Aspects of will included the overwhelming belief in the *value of PDSA* as reported by participants and a mixed response concerning *enthusiasm* or *frustrations* participants reported in their commitment to actually carry out PDSA. For capacity, features that participants pointed to as helping build capacity included *connections to previous experiences* and *PDSA training*, while challenges to capacity-building included *time* and *data collection and analysis*. In the following section, we first explore the will and capacity of innovation design team members for each district case, followed by a cross-district comparison.

Will in Broward

Value of PDSA. All of the district and school practitioners on the innovation design team in Broward (100%) reported that they saw the value in PDSA as an approach for improvement. They valued the results of data collection as a way to validate what they had already been doing, evaluate changes that needed to be made, and "see the process and actual product [innovation] working." Other participants reported the potential of PDSA to help students succeed, to "tweak" or "custom[ize]" the innovation to the needs of the school, to give teachers a voice, and to contribute to the larger work of implementing PASL at scale. This was corroborated by researchers, who further reported that, at times, the value that school team members placed on PDSA developed only after they tried it out. As one researcher reflected:

> Schools had balked at the idea of the amount of time that this might take, whether or not they were going to get a bunch of pushback from teachers. Then, when they went and did it, a lot of the teachers really liked it. A lot of the surveys they were getting back said,

> "Yeah, this is something that really helped me."… So, it was a really positive experience that I think really got people excited and willing to move forward.

Enthusiasm. Five of the 16 practitioners (31%) on the innovation design team in Broward expressed enthusiasm about actually implementing PDSA. They described PDSA as being "very useful" in helping the school to see how the innovation impacted achievement, to change things about the innovation to fit the needs of the school, and in preparation for broader implementation (i.e., to all 9th grade students). One school practitioner expressed excitement because everyone had a voice in the planning process, saying

> There were always different ideas and sometimes chaos is good because it means people are thinking and people have a vested interest in it. So, what went well was everybody had different ideas about what would work and what would not work, and we were able to finally discuss our differences…and come to an agreement on what will work.

Two of the district practitioners also indicated that they would continue to conduct PDSA in the future. Referring to using PDSA to solve other problems in the district, one said, "We're looking forward to where it goes from here." Two of the three researchers corroborated this enthusiasm, saying, "I saw them enthusiastic about trying something out, getting together to discuss what the results were, and then deciding what to do next."

Frustrations. Even though all practitioners noted the value of the idea of PDSA as an improvement approach and about one-third expressed enthusiasm in their commitment to implementing PDSA, six of the 16 district and school members (38%), one researcher, and one development specialist in Broward noted frustrations with implementing PDSA in the innovation schools. As a school design team member stated, "some people just won't do it." Four of the six specifically used the word "frustrating" to describe PDSA implementation and this frustration

stemmed from a few sources. Two felt "insulted" that PDSA was being taught as if it was something new to them. For example, a design team member explained, "We did some shoddy work on [PDSA], because we were just frustrated that we were having to go through the motions to pretend that this was something new."

Four practitioners mentioned pushback around completing forms from the facilitation guides, calling the process "laborious," "discouraging," and "redundant." As another school practitioner stated, "There was a time in the middle of this process, where the schools felt like we were trying to write somebody's doctorate paper." A researcher summarized this in saying, "the common phrase that we were hearing in Broward…was 'too much paperwork'…We don't want to do all this paperwork.' And so that became kind of the rallying flag in Broward against the idea of really doing PDSA cycles." Only one school practitioner expressed opposition to the general concept of continuous improvement, stating in a session feedback form: "We have the RCIs [established with] the RCI form. Why change [the process] with two months left in the school year? We need to continue to reinforce what we started instead of trying to change/add mid-stream." Of note, enthusiasm and frustrations were not mutually exclusive categories as three practitioners expressed both.

Will in Fort Worth

Value. Thirteen of the 15 school and district practitioners on the innovation design team (87%) articulated that they saw some value in going through the PDSA process. Five members, in particular, seemed to value the ability to "see what's working and what's not" and "tweak" something after a review. For example, one school practitioner described the strength of PDSA as "that there's a way to fix, that's not completely scratching out what you're doing, like you're not throwing it out, but you're refining it. You're making it better, and it's a spiral towards

whatever your goal is." Three practitioners saw value in PDSA as a way to be more "intentional" and to "target" or "hone in" on student needs and effective practices. One district practitioner saw the value of PDSA as a tool for scaling the innovation within the district: "It was a reasonable way to answer the question: how will a process scale from a small number of adoptees to a larger one over a series of iterations."

Enthusiasm. Eight of the 15 practitioners (53%) made statements that provided evidence of real excitement about the PDSA work that they were doing. They described being invested in the process and eager to "revisit and refine" what they were doing. It appeared that innovation team members' commitment to the innovation goals of student ownership contributed to their excitement about PDSA as a process. As one school practitioner said, "I think it's going to help our kids, and I think that the SOAR ideals of the PDSA cycle is a quality way to look at what you've done and really reach the kids in the way that you want to make them accountable for their own learning." More specifically, three members described how using PDSA facilitated bringing new personnel on board and "allowed SOAR to kind of take some roots." Showing buy-in and commitment to implementing PDSA beyond the SOAR innovation, one member reported using the PDSA process to work with a student teacher in his/her classroom. One researcher observed that enthusiasm for implementation increased over time, reporting that practitioners initially struggled but voiced support and enthusiasm for PDSA by the end of the school year.

Frustrations. Eight of the 15 practitioners (53%) made statements that reflected frustrations by members to the PDSA process. There was a notable school-level difference in this reported frustration. Of the eight members who expressed frustration, five of them worked at one school. As a development specialist noted about this school, "they'll make this bold proclamation that they're going to do the cycle and then [we] will try to call or email and it's

crickets chirping." This heightened frustration from the school, along with a general lack of support for CIR from the school's leadership, was evident during a cross-school meeting when the school team had nothing to present during time dedicated to reporting on the PDSA cycle. The three researchers had a starker perspective about practitioner frustrations with PDSA; one noted, "Nobody ever has bought into PDSA and it's been a struggle to get them to do it in all three schools."

Opposition to measuring outcomes during PDSA was evident across members from all three innovation schools during training sessions. Participant responses suggest that the primary cause of this was the accountability culture in the district. For example, during an introduction to PDSA early in the process, one practitioner noted, "I think everybody here is really on board and appreciates the work and likes what it looks like if we were just using it to study our model. But I don't know if any of us feel comfortable that this is all it would ever be used for." This suggests a fear among practitioners that the district would use PDSA data for purposes other than improvement. Further, some practitioners referenced a district policy specifying that a beneficial program should not be artificially withheld from some students for research purposes as conflicting with the idea of using PDSA to test changes on a small scale. As in Broward, frustrations and opposition were not mutually exclusive with enthusiasm, as seven practitioners expressed both.

Will Cross-District Comparison

In both districts, participants believed in the value of PDSA but demonstrated mixed feelings about whether it was worth putting in the time and effort necessary for successful implementation. For many, their belief in the value of PDSA reflected the intended purposes of improvement science: identify what works and what doesn't, tweak an innovation to fit their

local context, leverage effective teacher practices and student success, and give practitioners a voice in reform. Although there was buy-in for the idea of PDSA, participants' desire to actually implement PDSA was mixed. About half of the practitioners in each district reported enthusiasm toward implementation, while about half reported frustrations.

Participants in the two districts differed in their reasons for opposing PDSA; these differences appeared to be due to contextual factors. In Broward, participants were frustrated by the amount of paperwork they had to complete as it seemed to go against what they understood PDSA to entail based on their prior experiences. In Fort Worth, the district culture of accountability deterred some from buying in to PDSA. Practitioners thought that the district would not value data other than test scores or would use the data collected to hold them accountable. Furthermore, practitioners from one of the three schools in Fort Worth expressed more frustrations regarding the SOAR innovation than practitioners in the other schools, suggesting an influence of school-level contextual factors.

Capacity in Broward

Connections to previous experiences. Six innovation design team members (38%) mentioned the connection between PASL PDSA cycles and previous experiences with continuous improvement in the district. Five participants agreed that previous exposure to PDSA and continuous improvement made the process of implementation easier because it "kind of hit home."

In contrast, one participant noted that the "minutia" of the PASL PDSA cycles was different from previous district applications of PDSA cycles, and further noted that the lack of any acknowledgment by the trainers that this was something that the school teams were already

familiar with was frustrating. A district representative concurred. In describing what he would have changed about the PDSA training, he stated,

> If you had a perfect world…you would also say, "We know you have done this, we know where you are with the continuous improvement plan in the state of Florida, and we're going to just give you a little greater information on what we're going to be looking at in a shorter cycle period."

These two comments suggest that had explicit discussion of similarities and differences between practitioners' prior PDSA experiences and the PASL PDSA process been made, greater capacity could have been built.

PDSA training. Feedback collected at the end of training sessions that covered PDSA indicated that all but two innovation team members (88%) felt prepared to carry out tests in their schools. In fact, PDSA trainings included features that the professional development literature has shown to be associated with teachers' willingness and capacity to change their practice, including sustained and scaffolded trainings over time, hands-on participation that engages implementers in authentic experiences, and the use of aligned materials that provide scaffolding for implementers' learning (Borko, 2004; Cohen & Hill, 2002; Garet et al., 2001).

However, three participants did not perceive this to be the case. One participant perceived the trainings, particularly the webinars, to be "busy work," reporting,

> It felt a little bit like busy work because there was a guided template and we were being facilitated through the process, two and a half hours of reflecting and discussing and brainstorming and what have you. But, at the end of the day, that really wasn't helpful as to our individual campus trying to address our own concerns and facilitate the pilot on our end.

This suggests that facilitating the PDSA process across three different school sites may have negatively impacted the degree to which the training sessions were connected in authentic ways to participants' local contexts.

Additionally, the trainers (researchers and development specialists) reported that they were still building capacity in themselves. Members of the research team expressed that they were not always clear what role they should take in the PDSA process, and this may have impacted their ability to build their own knowledge and capacity on the appropriate way to support implementers in the process. One researcher acknowledged, "We're kind of in a new territory." Another researcher suggested that the development specialists struggled with the role in facilitating PDSA, saying,

> We assumed that developers would have some capacity to do this. Unfortunately it turns out that they didn't, and neither did any of us… So, we're trying to sit down and put together a manual for implementing PDSA in the district, and [at the same time] we're going down and trying to build other people's capacity for doing PDSA in a district.

A development specialist similarly expressed some challenges around the capacity to facilitate CIR work, including not "being on the same page" with researchers in regard to their understanding of PDSA.

Time challenges. Of the various issues related to the capacity to implement PDSA, time was the most prominent challenge that practitioners reported. Ten of 16 Broward practitioners (63%), in addition to two development specialists and researchers noted that school staff, particularly teachers, did not have sufficient time to engage fully in data collection and analysis activities associated with the "Do" and "Study" phases. In particular, participants indicated that the paperwork and monitoring required of the data collection process was burdensome for

teachers, particularly given pressures such as standardized testing schedules and curriculum requirements. For example, one practitioner noted,

> [It's a] great process if you're not so bogged down with testing and all the other elements that comes with, you know, teaching a curriculum across the year. [...] There's very little room to do the cycle more than once in a given semester.

Similarly, another practitioner noted that PDSA, particularly filling out templates provided in the PDSA facilitation guide, was viewed as an extra thing that school personnel "don't have time to write." Two practitioners further mentioned that the timing of introducing PDSA in the middle of the school year hindered successful implementation.

Researchers and development specialists also expressed concerns about time. The geographic distance between the PDSA trainers and the school teams posed a logistical challenge as it constrained how often they could schedule time to meet with the school teams in person. While webinars were often used in place of face-to-face meetings, a researcher and two design team members concurred that interactive processes like 'Study' meetings did not work well through webinars.

Data collection and analysis challenges. The ability to regularly collect and analyze data was also perceived to be a second central challenge of PDSA implementation. This issue was closely related to time and resources, as practitioners described the data collection process as burdensome and potentially overwhelming for teachers. Additionally, in spite of efforts by development specialists and researchers to explain the importance of data in the PDSA process, several school practitioners were confused about the purpose of collecting and analyzing data. One practitioner noted, for example,

> We were doing these sort of small steps of collecting data, but what was the purpose of us collecting the data? Was it to see if the method of data collection was appropriate or was it to see if there was change in the data? I don't think that was always clear with everybody.

Researchers and practitioners further reflected that design team members seemed unclear about their role in the regular collection and analysis of data and how data should be used. While data collected in PDSA cycles is meant to be used for refining an innovation (Lewis, 2015), a handful of school practitioners only noted the usefulness of data for realizing the distal goals of the PASL innovation.

Capacity in Fort Worth

Connections to previous experiences. With the exception of the district central office personnel on the innovation design team who had experience with PDSA from the district's Baldridge approach, only one practitioner mentioned having some prior exposure to PDSA. Though lacking previous experience with an explicit PDSA model, two-thirds (66%) of design team members connected the PDSA process to their everyday practices, such as reflecting on their lessons or using the scientific method. Representative practitioner comments included "it's something that people kind of do naturally" and "I think [PDSA is] just a good tool to use with any teacher in any lessons is to see if that lesson worked and go back and do what would be better for the kids' learning." A science teacher similarly noted, "In my mind, it's just a scientific method...it just helps guide your thinking and your planning and, most importantly, the next steps."

Amidst these widespread perceptions that PDSA was a "natural" or everyday practice, researchers did note two key challenges that stemmed from practitioners' connections to

previous experiences. First, teachers were familiar with using district or state assessment data, but had less experience designing a data collection strategy. One researcher noted, "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." The second key challenge this researcher noted was the difficulty in changing practitioners' mindsets from determining an implementation plan for an entire year, a familiar practice for Fort Worth practitioners, to testing one discrete change idea. The researcher explained,

> They wanted to kind of come up with their overall plan. 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,' and we were trying to say to them, 'well, how do you know you want to do something unless you know it works?'

PDSA training. Feedback after training sessions indicated that the sessions helped practitioners see the utility of PDSA to refine the innovation and come up with a working plan for the school. But, as in Broward County, innovation design team members also noted some challenges during the training. These challenges included ambiguity in the goals of the project, unclear expectations, and participants finding the trainings "too bookish" and disconnected from the work they were doing in their schools.

In addition, development specialists and researchers acknowledged the challenges of trying to build capacity in practitioners for a process that they themselves were still learning. They noted that the development specialists, in particular, were not well prepared to support some of the technical details of PDSA occurring between training sessions. Consequently, one researcher explained that practitioners were allowed to have "less precise … PDSA work than

might have been helpful." Though development specialists indicated that they still had much to learn, they remained optimistic. As one development specialist explained:

> We're new to learning the process and while it seems very powerful, it's very complex and we're operating at somewhat of a deficit in that we haven't had a lot of practice in – and training, so we're going to – we're committed and we're going to work hard at it, but the [participants'] capacity to do this work is going to be directly connected to their ability to get the skills that they need to do it.

Despite these challenges, several design team members expressed optimism that PDSA would be successful the next year. Practitioners felt they had clearer expectations and learned about the process. Development specialists also felt better prepared. As one explained, "I am very optimistic. I'm actually really quite enthusiastic and so I think the more we know, the more the increased likelihood that the [innovation design team] will have that skill and capacity."

Time challenges. When asked to identify the main challenges with PDSA, two-thirds (66%) of design team members, all three development specialists, and two of the three researchers mentioned issues of time. Finding enough time in the day given their roles as full-time teachers, coaches, department chairs, and other commitments was a tremendous difficulty. One school member explained, "The challenges were definitely the time to fit in the meetings on campus…to collect the data, to have that continuous conversation between the team members, because everyone on the team at our school is highly involved." Practitioners were concerned about finding time not only within their busy schedules, but within their curricular calendars. Despite some relief that the initial lesson development was done for them by the development specialists, some practitioners expressed concern that piloting the lessons on growth mindset and

problem solving would take too much time away from delivering the district's curriculum to students.

In addition to finding time for PDSA within their busy days, practitioners also struggled with the timing of when PDSA was introduced, which coincided with the spring testing calendar. One design team member noted, "at the point where we were implementing, we had a lot of state testing starting and AP, and just different testing and the school winding down....I think if we would have started the cycles at the beginning of the school year, I think we'd have seen even more involvement from different teachers."

Data collection and analysis challenges. A second major challenge in implementing PDSA in Fort Worth was collecting and analyzing data, mentioned by eight of the fifteen innovation design team members (53%), all three development specialists, and two of the researchers in Fort Worth. In addition to the difficulty of finding time to collect and analyze data, the challenges were two-fold. First, participants struggled with understanding the value of the systematic collection of data, rather than relying on informal reflection. One practitioner summarized the data collection challenge by saying, "We just plan, do, and kind of -- oh, that kind of worked, oh, that kind of didn't work... I think that the Study [phase] is the weakest for teachers." Similarly, a researcher noted that "it was sometimes a struggle to get schools to understand the value of doing much more than kind of say, you know, testing something and then saying, 'hey, how'd it go?'"

The second data challenge was determining what data the innovation design teams wanted to collect and designing appropriate data collection instruments. Participants were inexperienced in this type of activity and expressed that they felt uncomfortable designing

questionnaires. In addition, they lacked clarity over the desired outcomes for which they were supposed to collect data. As one practitioner shared,

> We were creating a teacher feedback form to collect data. We talked about what type of questions to put on the questionnaire over the course of a week. I think we were all just discussing it, either in person or through email, because there's a little bit of ambiguity about how to collect the specific data we needed, and what kind of questions would get that data.

A researcher similarity noted the continuing challenge of designing new data collection instruments that would not be too burdensome to administer.

Compounding the struggles members had with data collection was confusion around who was responsible for the "Study" phase of PDSA. Some design team members thought that the researchers were the "studiers" and, indeed, one researcher told the school team she was working with that the research team would take over the compiling and analyzing of data. Similarly, one practitioner referred to "the study people" as the external researchers and expressed a desire that "the study people could find the main things that were effective from those other schools...and just present that to us."

Capacity Cross-District Comparison

During cross-district webinars and on session feedback forms, innovation design team members in both districts overwhelmingly reported a high degree of confidence in their ability to conduct PDSA. In both districts, this confidence in capacity to implement PDSA cycles could be attributed to the connections that practitioners were able to make between PDSA and previous experiences (e.g., other continuous improvement efforts, reflective teaching practices) and the PDSA training led by the development specialist and researcher partners. Yet a handful of

participants in both districts noted that the "we do this already" mentality became an obstacle to deeper change as a handful of practitioners did not adjust their preconceived notions of PDSA, reflective practice, and data collection. Likewise, researchers and development specialists who conducted the PDSA trainings expressed that they had not yet built PDSA capacity in themselves, resulting in training that a number of participants across both districts perceived as ambiguous or disconnected from school practices.

At the same time, interviews and design team member reflections from both districts uncovered challenges in building capacity to implement PDSA pertaining to time and data. Across both districts, practitioners found that with the data collection expectations and tools the researchers provided, they struggled to find enough time in the day to meet, collect data, and fill out paperwork to document each cycle. They also struggled with the timing of PDSA's introduction in the middle of the school year, which hindered each school's ability to conduct multiple cycles. Regarding data collection and analysis, participants in both districts struggled with understanding the purpose of data collection during the PDSA cycle and their responsibilities in the process.

There were nuanced differences in the ways that these broader capacity supports and challenges manifested themselves in each district. In Broward, participants made connections to the district's PDSA strategy and wished that the trainers would have made explicit connections to the district's prior CIR work. In contrast, Fort Worth innovation team members tended to draw connections between PDSA and specific teaching practices common to their subject area, for example, the scientific method. In addition, Fort Worth participants alluded to multiple competing priorities that limited their time and ability to conduct PDSA cycles, including the district curriculum, the school testing calendar, and plans to implement the innovation for the

whole school. Just as PDSA is a process used to adapt innovations to local contexts, local contextual factors such as prior experiences, competing policy demands, and professional norms of practice influenced practitioners' perceived capacity to implement PDSA.

Discussion

In unpacking the perceived will and capacity of high school practitioners to engage in PDSA cycles for the first time in the first year of implementation, these findings reveal possibilities and challenges in conducting continuous improvement research in high schools. First, across both districts, participants' perceptions of will revealed that school and district practitioners saw value in PDSA. This is promising, as belief can be used to motivate individuals to commit energy to implementing a new practice (McLaughlin, 1990). Findings also corroborated McLaughlin's (1990) proposition that belief can follow practice; practitioners and researchers noted that, for some, enthusiasm came after they had begun "practicing" PDSA at their respective school sites. Initial motivation, therefore, may not need to be present for PDSA implementation to occur. Once practitioners begin conducting PDSA cycles, they may recognize the value of the approach and take ownership of it as a method for improvement in the future.

A second promising finding for the uptake of PDSA in high schools was that it built on practices that participants reported they were already doing. While this perceived familiarity did pose some sensemaking challenges, described below, it also suggests that PDSA may be an incremental rather than radical change to current school practices in some districts. Compared to the implementation of radical change, carrying out incremental change has been found to generate less resistance among implementers (Tyack & Cuban, 1995).

Yet these findings also reveal a number of challenges faced by teams of practitioners, researchers, and development specialists engaged in PDSA cycles in high school contexts. These

challenges suggest three important shifts that, as CIR is increasingly incorporated into education (Identifying Reference, 2015), organizations leading such efforts need to address. One challenge our data illuminated is a contradiction between practitioners' perceptions of the idea of PDSA and whether they had the motivation, desire, and ability to actually carry it out. In short, practitioners embraced the idea of PDSA and thought it was similar to what they already do, but also felt it was disconnected from their daily work. One explanation for this apparent contradiction is that as practitioners made sense of PDSA, their initial perception that it was similar to a more general process of reflection may have inhibited their conceptual understanding of what is unique about PDSA. A cognitive theory of implementation (Spillane et al., 2002) explicates that individuals interpret new information in light of the pre-existing schemas they already hold: co-opting new ideas to fit with what they already know, grasping onto surface-level similarities while ignoring differences that require deeper change, and preserving one's self-image and self-esteem by discounting reforms that suggest that what one was doing before was "wrong" or "less effective than what we could be doing" (p. 402). Evidence of this could be seen here in the way that design team members expressed frustration that their previous experiences with PDSA were not acknowledged.

The implications of this for those engaged in training practitioners in PDSA suggest that the learning opportunities provided to design team members would have been improved by introducing PDSA in ways that did not make them feel like their prior knowledge and practices were being ignored or undermined, but also that the new process would require more of them than had previously been demanded in their districts. In particular, trainers should introduce PDSA not as something entirely new, but rather as something that builds on practitioners' prior knowledge and current practices (e.g., explaining that PDSA adds a layer of formality to

reflective practice by having teachers document the impact of a lesson and come up with an action plan based on their findings). This would require that those facilitating PDSA trainings systematically examine what practitioners' prior experiences with PDSA or similar practices look like. To do so, trainers might ask practitioners to model those practices, and then engage in critical conversations that point out how PDSA might differ.

With this in mind, trainers should also attend to how educators use different types of data, such as state assessments, district benchmarks, and classroom assessments, in different ways (Farrell & Marsh, 2016), with the goal of exploring the attributes of data they need to collect in a PDSA cycle to answer their current questions (Supovitz, 2012). Prior research has highlighted that, to influence practice, teachers need fewer but quicker data collection tools, but that producing these tools and interpreting the resultant data requires substantial expertise (Lynch, Smith, Provost & Madden, 2016). This reflects our findings on the challenge of building capacity in skills to collect and analyze data. To be successful, trainers need to provide ongoing technical support in this area (Marsh, 2012). That being said, trainers should plan their scope and sequence of ongoing technical support with the understanding that these skills take time to development. As our data only captured the first year of PDSA implementation, perhaps one year is not enough to build full capacity. Instead, researchers may need to scaffold practitioners' responsibility for data collection and analysis over multiple years.

A second important shift that CIR advocates need to consider comes from our finding that there was a lack of joint ownership over PDSA processes between the researchers, development specialists, and practitioners. Even as practitioners valued PDSA and the innovations they were testing through PDSA, they resisted the specific forms they had to fill out for each phase and the scheduling of when the cycles would take place. At times, this caused

frustration amongst practitioners over the tasks they were being asked to complete. We posit that this challenge comes from asking practitioners to take on roles as researchers, which required a shift in their mindset about the nature of their work. Indeed, teachers are not accustomed to thinking of their classroom lessons as offering knowledge for others outside of their context (Hiebert, Morris, & Glass, 2003), yet that is what the collective PDSA process assumes. CIR asks practitioners and researchers to serve as boundary crossers, to use the tools and objectives of other types of professionals and integrate them into their work (Schenke, van Driel, Geijsel, & Volman, 2017). Yet, teachers do not see the iterative cycle of improvement as meeting their needs as they need to move on once a lesson is tried out and improved (Morris, Hiebert, Sisofo, & Hwang, 2015).

There may be some ways that CIR advocates can address the challenge this mindset shift requires, such as engaging with practitioners from the beginning in co-designing how the improvement process should be carried out in specific school sites. This may have potential for creating more authentic connections to training sessions and teachers' lived classroom experiences, an element of effective professional development concerning changes in teacher practices (e.g., Borko, 2004). In this regard, PDSA trainers could learn lessons from participatory action research, wherein "facilitators" (e.g., university professors) are not viewed as neutral researchers but as a part of the collective research group that co-constructs knowledge at every phase of the research and reflection process (Kemmis & McTaggart, 2005). Yet, if resistance to this shift continues, it may underscore the need for a larger reorganization as CIR is taken up by educators.

The third challenge highlighted by our findings, the difficulties of time and expertise for PDSA work, also suggests a need for a new type of infrastructure to support CIR. This study

identified pockets of frustration which stemmed from a sense that PDSA was too much work, and participants reported both a lack of time and confusion about the purpose of and how to collect and analyze data within an improvement science framework. These challenges are not surprising given the multiple policy demands, many of which incorporate data use, that school practitioners must balance (Hatch, 2001; Identifying Reference, 2011) and the newness of applying CIR to education Furthermore, they reflect similar challenges found among healthcare practitioners implementing PDSA, including issues around data collection and use (Reed, 2015; Taylor et al., 2014) and not having adequate time to implement (Berwick, 1998). The implications for those conducting PDSA work suggests that they should engage school and district administrators in establishing organizational structures within the school to facilitate improvement work. Indeed, substantial research on educator data use indicates that the organizational structures and culture around data use initiatives shape how individuals use and interpret data (Honig & Venkateswaren, 2012; Marsh, 2012). Our findings indicate that finding time to engage in PDSA data collection and analysis work hampered educator capacity for PDSA. This is consistent with prior research on the importance of providing organizational structures that allow time for collaboration among stakeholders to examine evidence and come to collective interpretation (Marsh, 2012). The organizational culture around the purposes of data use, for accountability or for improvement, also shape how educators respond to such initiatives (Cannata, Redding, & Rubin, 2016; Weiss, 2012). Building capacity for PDSA requires helping educators understand how using data for improvement, which is the goal of PDSA, differs from using data for accountability.

More generally, this study points to the need to develop to develop the improvement infrastructure to support CIR. While educational research has greatly expanded the infrastructure

to support rigor research methods, there is little infrastructure to support those engaged in improvement, particularly continuous improvement efforts (Peurach, 2016). Yet school, district, and system infrastructure can support or inhibit collaborative learning to improve practice (Hopkins & Woulfin, 2015). CIR must contend with the nature of the infrastructure to support PDSA in education, such as how involved practitioners themselves are in PDSA and how much of the measurement and analytic infrastructure is led by individuals in the network hub (Russell et al., 2017).

As the first study to examine educators' will and capacity to implement PDSA, our work has generated further questions and additional areas of future inquiry into the implementation of continuous improvement processes. First, because the present study only examined will and capacity of novice innovation team members during the first year of implementation, future researchers might ask whether will and capacity changes over time as CIR participants gain more experience in training and implementation. Second, future research might examine how the nature of the innovation (i.e., what is being tested) may condition will and capacity, exploring whether innovations with certain characteristics (e.g., lesson-based vs. culture-building) are better suited than others to continuous improvement processes in high school contexts. While in the present study, practitioners did not choose the focus of the innovation (i.e., PASL and SOAR), future research in this line of inquiry might also examine if practitioner decision-making in the innovation to develop and refine may impact their will to conduct PDSA cycles around it.

A third avenue for future research could expand upon a finding related to the lack of initial capacity amongst the researchers and development specialists who served as first-time PDSA trainers. Our data sources were limited in that they did not explicitly ask for information related to the will and capacity of researchers, rather than practitioners, to implement PDSA.

Continuous improvement research involves a shift in the role of the researcher to one of collaborative partner, and future research should explore in greater depth the processes through which researchers and other external partners develop their own capacity and desire to adapt to this new role.

Fourth, future research should link educators' will and capacity to implement PDSA to implementation outcomes, in this case whether practitioners tweaked the respective PASL and SOAR innovations based on the data they collected during PDSA cycles. Furthermore, researchers should examine whether implementing PDSA cycles to tweak PASL and SOAR innovations lead to desired student-level outcomes, including attendance, course passing rates, and student discipline.

Conclusion

Our research contributes new understandings of specific factors that may thwart or support high school practitioners' initial will and capacity to implement improvement work that would ultimately lead to improved student outcomes. As CIR continues to gain traction in the field of education, we invite researchers to follow through on our recommendations for future research and "test out" our findings in other school contexts.

References

Anderson, T., & Shattuck, J. (2012). Design-based research a decade of progress in education research? *Educational Researcher,* 41, 16-25.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences, 13*(1), 1-14.

Berwick, D. M. (1998). Developing and Testing Changes in Delivery of Care. Ann Intern Med. 128:651-656

Bhuiyan, N., & Baghel, A. (2005). An overview of continuous improvement: From the past to the present. *Management Decision, 43*(5), 761-771.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8), 3-15.

Bryk, A.S. (2009). Support a science of performance improvement." *Phi Delta Kappan, 90*(8), 597- 600.

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

Cannata, M., Redding, C., & Rubin, M. (2016). Continuous Improvement in Action: Educators' Evidence Use for School Improvement. Presented at the annual meeting of the Assocation for Education Finance and Policy, Denver, CO.

Cobb, P., Jackson, K., Smith, T., Sorum, M., & Henrick, E. (2013). Design research with educational systems: Investigating and supporting improvements in the quality of mathematics teaching and learning at scale." In *Design Based Implementation Research: Theories, Methods, and Exemplars*, B.J. Fishman, W.R. Penuel, A.R. Allen, and B.H. Chang. New York: Teachers College.

Cohen, D.K., & Hill, H.C. (2001). *Learning policy: When state education reform works*. New Haven: Yale University Press.

Deming, W.E. (2000). *The new economics for industry, government, and education*. Cambridge, MA: The MIT Press.

Devos, G., Tuytens, M., & Hulpia, H. (2014). Teachers' organizational commitment: Examining the mediating effects of distributed leadership." *American Journal of Education, 120*(2), 205 – 231.

Dolle, J. R., Gomez, L. M., Russell, J. L., & Bryk, A. S. (2013). More than a network: Building professional communities for educational improvement. *National Society for the Study of Education, 112*(2), 443-463.

Farrell, C. C., & Marsh, J. A. (2016). Metrics Matter How Properties and Perceptions of Data Shape Teachers' Instructional Responses. *Educational Administration Quarterly*, *52*(3), 423–462. http://doi.org/10.1177/0013161X16638429

Fishman, B.J., Penuel, W.R., Allen, A., & Cheng, B. H. (Eds.). 2013. *Design-based implementation research: Theories, methods, and exemplars. National Society for the Study of Education Yearbook*, Volume 112, Issue 2 (pp. 136 - 156). New York, NY: Teachers College Record.

Firestone, W.A. (1989). Using reform: Conceptualizing district initiative. *Educational Evaluation and Policy Analysis, 11*(2), 151-164.

Garet, M.S., Porter, A.C., Desimone, L., Birman, B.F., & Suk Yoon, K. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*(4), 915-945.

Hiebert, J., Morris, A. K., & Glass, B. (2003). Learning to Learn to Teach: An ``Experiment'' Model for Teaching and Teacher Preparation in Mathematics. *Journal of Mathematics Teacher Education*, *6*(3), 201–222. https://doi.org/10.1023/A:1025162108648

Honig, M. I., & Venkateswaran, N. (2012). School–Central Office Relationships in Evidence Use: Understanding Evidence Use as a Systems Problem. *American Journal of Education*, *118*(2), 199–222.

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

Kemmis. S. & McTaggart, R. (2005). Participatory action research: Communicative action and the public sphere. In N.K. Denzin, & Y.S. Lincoln (Eds.). *The Sage handbook of qualitative research* (pp. 559 - 603). Thousand Oaks, CA: Sage Publications.

Langley, G.J., Nolan, K.M., Norman, C.L., & Provost, L.P. (2009). *The improvement guide: practical approach to enhancing organizational performance* (2nd ed.). San Francisco, CA: Jossey Bass.

Leithwood, K., and Jantzi, D. (2000). The effects of transformational leadership on organizational conditions and student engagement with school. *Journal of Educational Administration, 38*(2), 112-129.

Lewis, C. (2015). What is improvement science? Do we need it in education? *Educational Researcher, 44*(1), 54 – 61.

Lynch, D., Smith, R., Provost, S., & Madden. (2016). Improving teaching capacity to increase student achievement: The key role of data interpretation by school leaders. *Journal of Educational Administration*, *54*(5), 575–592. http://doi.org/10.1108/JEA-10-2015-0092

Mann, N.R. (1993). *Statisticians in history: W. Edwards Deming*. Retrieved from http://www.amstat.org/about/statisticiansinhistory/index.cfm?fuseaction=biosinfo&BioID

Marsh, J. A. (2012). Interventions Promoting Educators' Use of Data: Research Insights and Gaps. *Teachers College Record*, *114*(11), 110309.

Marshall, C., & Rossman, G.B. (2011). *Designing Qualitative Research (5th edition)*. Thousand Oaks, CA: Sage Publications.

McLaughlin, M.W. (1990). The RAND change agent study revisited: Macro perspectives and micro realities. *Educational Researcher, 19,* 11-16.

Means, B., & Harris, C. J. (2013). Towards an evidence framework for design-based implementation research. In B.J. Fishman, W.R. Penuel, A.R. Allen, and B.H. Chang (Eds.), *Design based implementation research: Theories, methods, and exemplars*. New York, NY: Teachers College,

Miles, M., & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage Publications.

Morris, A.K., & Hiebert, J. (2011). Creating shared instructional products an alternative approach to improving teaching. *Educational Researcher, 40*(1), 5-14.

Morris, A., Hiebert, J., Sisofo, E., & Hwang, S. (2015). Using evidence in classroom practice. Presented at the National Center on Scaling Up Effective Schools Conference: Using Continuous Improvement to Integrate Design, Implementation, and Scale, Nashville, TN: Vanderbilt University.

Orginc, G. & Shojania, K.G. (2014). Building knowledge, asking questions. *BMJ Quality and Safety, 23*, 265-267.

Park, S., Hironaka, S., Carver, P., & Nordstrum, L. (2013). *Continuous improvement in education*. Palo Alto, CA: Carnegie Foundation for the Advancement of Teaching.

Penuel, W. R., Allen, A., Coburn, C. E., & Farrell, C. (2015). Conceptualizing research–practice partnerships as joint work at boundaries. *Journal of Education for Students Placed at Risk, 20*(1-2), 182-197.

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

Russell, J. L., Bryk, A. S., Dolle, J. R., Gomez, L. M., LeMahieu, P. G., & Grunow, A. (2017). A Framework for the Initiation of Networked Improvement Communities. *Teachers College Record*, *119*(7).

Sanders, M.G. (2014). Principal leadership for school, family, and community partnerships: The role of a systems approach to reform implementation. *American Journal of Education*, *120*(2), 233 – 255.

Shortell, S.M., Bennett, C.L., & Byck, G.R. (1998). Assessing the impact of continuous quality improvement on clinical practice: what it will take to accelerate progress. *Milbank Quarterly, 76*(4), 593-624.

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.

Sowers, N., & Yamada, H. (2015). *Pathways impact report*. Palo Alto, CA: Carnegie Foundation for the Advancement of Teaching.

Stein, M.K, & Coburn, C.E. (2007). Architectures for learning: A comparative analysis of two urban school districts. *American Journal of Education, 114*(4), 583-626.

Supovitz, J. A. (2012). Getting at Student Understanding—The Key to Teachers' Use of Test Data. *Teachers College Record*, *114*(11), 110309.

Taylor, M.J., McNicholas, C., Nicolay, C., Darzi, A., Bell, D., & Reed, J.E. (2014). Systematic review of the application of the plan-do-study-act method to improve quality in healthcare. *BMJ Quality and Safety, 23*, 290-298.

Tseng, V. (2012). Forging common ground: Fostering the conditions for evidence use. William T. Grant Foundation Annual Report.

Tyack, D. & Cuban, L. (1995). *Tinkering toward utopia: A century of public school reform*. Cambridge, MA: Harvard University Press.

Weatherley, R. & Lipsky, M. (1977). Street-level bureaucrats and institutional innovation: Implementing special education reform. *Harvard Educational Review, 47*(2), 171 – 197.

Weiss, J. A. (2012). Data for Improvement, Data for Accountability. *Teachers College Record*, *114*(11), 110309.

Yin, R.K. (2009). *Case study research: Design and methods* (Vol. 5). Thousand Oaks, CA: Sage Publications.

Table 1. District Characteristics

	Broward County Public Schools (FL)	Fort Worth Independent School District (TX)
District context		
Student enrollment	262,563	84,588
Student demographics	40% Black, 30% Hispanic, 24% White 60% economically disadvantaged 14% English Language Learner	63% Hispanic, 23% Black, 11% White 77% economically disadvantaged 31% English Language Learner
Number of high schools	33	16
State accountability context	A Plus (A-F) school grading system introduced in 1999; high schools assigned grades based on student performance, participation/performance in accelerated curricula, graduation rate, and college readiness. High-stakes testing for students and teachers Teacher evaluation includes value-added measures.	First school accountability ratings introduced in 1993. New high school assessment system (a set of twelve End of Course exams) introduced in 2011-12. High schools rated as "Met Standards" or "Improvement Required," based on meeting overall achievement benchmark, closing performance gaps, and postsecondary readiness measures.
Continuous improvement experience	2012-2013 strategic plan included continuous improvement as one of three core goals. Limited district capacity to design new programs, so they adopt established programs	2013 strategic plan adopted Baldridge Continuous Improvement Cycle to monitor and improve district performance. This includes PDSA cycles. In 2013-14, PDSA evident mostly in district central office operations.
District innovation design team composition		
Development specialists	2	3
Researchers	3	3
Teachers/faculty	10	9
School administrators	2	5
District administrators	3	5
School innovation design team composition		
Teachers/faculty	11	14
School administrators	2	1

	Broward County Public Schools (FL)	Fort Worth Independent School District (TX)
Innovation characteristics		
Name	Personalization for Academic and Social Emotional Learning (PASL)	Student Ownership and Responsibility (SOAR)
Major components	1. Assign each 9th grade student to a PASL teacher 2. Form educator teams to respond to students' individual needs 3. Make intentional points of contact (IPCs) with students through Rapid Check-Ins (RCIs) and Problem-solving meetings (PSMs) 4. Create norms for engagement to build a PASL culture 5. Provide instruction in goal achievement skills 6. Make intentional use of information to plan PASL activities and identify students who need additional support	1. Develop growth mindsets in students through introductory lessons and classroom extension practices 2. Build student problem-solving skills through explicit instruction and application in academic and social-emotional challenges 3. Goal setting and planning instruction to integrate problem-solving in academic context 4. Behavioral reflection sheet that encourages students to reframe their mindset and work through problem-solving steps 5. Teacher professional development and on-going learning in Professional Learning Communities

Table 2.

District Comparison of Will and Capacity to Implement PDSA Cycles

	Broward County Public Schools	Fort Worth Independent School District
Will	Value of PDSA (100%) Enthusiasm (31%) Frustrations (38%) - Paperwork	Value of PDSA (87%) Enthusiasm (53%) Frustrations (53%) - Accountability - School context
Capacity	Connections to previous experience (38%) - PDSA work PDSA training Time challenges (63%) - Meetings - Paperwork Data collection and analysis challenges	Connections to previous experience (66%) - Similar teaching practices PDSA training Time challenges (66%) - Meetings - Paperwork - District calendar Data collection and analysis challenges

Appendix A – Sample Form from PDSA Facilitation Guide

Form 2: PDSA Cycle Form

Tester Name(s):
Test Title (Brief):
Date(s) of Test: **Cycle #:**

PLAN (Document what you need to do before you run the test.)

Test Aim: What are we trying to accomplish?

Predictions: What do we predict will happen as a result of this practice?

1.
2.
3.
4.

Details: Finalize the details of the test, decide how to collect data, and assign individual school team members clear roles and responsibilities.

Test Details	Data Collection Details