

Chapter 14

Research on Continuous Improvement: Exploring the Complexities of Managing Educational Change

MAXWELL M. YURKOFSKY

AMELIA J. PETERSON 

JAL D. MEHTA

REBECCA HORWITZ-WILLIS

KIM M. FRUMIN

Harvard Graduate School of Education

As a result of the frustration with the dominant “What Works” paradigm of large-scale research-based improvement, practitioners, researchers, foundations, and policymakers are increasingly embracing a set of ideas and practices that can be collectively labeled continuous improvement (CI) methods. This chapter provides a comparative review of these methods, paying particular attention to CI methods’ intellectual influences, theories of action, and affordances and challenges in practice. We first map out and explore the shared intellectual forebears that CI methods draw on. We then discuss three kinds of complexity to which CI methods explicitly attend—ambiguity, variability, and interdependence—and how CI methods seek a balance of local and formal knowledge in response to this complexity. We go on to argue that CI methods are generally less attentive to the relational and political dimensions of educational change and that this leads to challenges in practice. We conclude by considering CI methods’ aspirations for impact at scale, and offer a number of recommendations to inform future research and practice.

We are in the midst of an exciting shift in education research and practice. As a result of increasing frustration with the dominant “What Works” paradigm of large-scale research-based improvement (Bryk et al., 2015; Penuel et al., 2011), practitioners, researchers, foundations, and policymakers are beginning to favor good

Review of Research in Education

March 2020, Vol. 44, pp. 403–433

DOI: 10.3102/0091732X20907363

Chapter reuse guidelines: sagepub.com/journals-permissions

© 2020 AERA. <http://rre.aera.net>

practice over best practice, local proofs over experimental evidence, adaptation over faithful implementation, and a focus on practitioners' problems over researchers' solutions. These ideas are embodied in a number of educational improvement methods that range in their origin and theory of action but are increasingly being labeled *continuous improvement* (CI) methods.

On initial inspection, CI methods have important differences. Some emerge out of the research–practice partnership (RPP) and design-based research traditions, therefore placing greater value on the role of researchers in supporting larger scale and multiyear research and improvement efforts. Others have developed from a focus on data-based decision making and professional learning communities, and therefore see smaller teams of educators as the core drivers of research and improvement. But underneath these differences, CI methods share a number of common characteristics that make them useful to analyze as a group. In particular, we are interested in four shared commitments:

1. Grounding improvement efforts in local problems or needs
2. Empowering practitioners to take an active role in research and improvement
3. Engaging in iteration, which involves a cyclical process of action, assessment, reflection, and adjustment
4. Striving to spur change across schools and systems, not just individual classrooms

These similarities also reflect the fact that, as illustrated in this chapter, CI methods share a similar intellectual lineage, one rooted in John Dewey's pragmatism and often connected to theories of organizational learning, quality improvement, action science, improvement science, design-based research, and teacher research.

In light of their alignment around these four pillars and their divergence across a number of other dimensions, we believe there is much to be learned by putting different CI methods into conversation with one another. With a few important exceptions that we discuss later (e.g., Coburn et al., 2013; Lochmiller & Lester, 2017; Peurach et al., 2018), much of the writing about these methods has (a) been written by proponents, (b) examined a single method only, or (c) emphasized the similarities among the methods as part of an effort to build a case for these approaches as a whole (Penuel et al., 2018). We offer an external examination of the methods in relation to each other and to their intellectual forebears, as well as the empirical evidence about their successes and challenges in practice.

We approach this task as a critical friend of the movement, sharing many of its aspirations but also seeking to clearly see the challenges, all in the hope of helping those who use these methods land in a better place. If we had to capture our main message in a sentence, it would be that CI methods as a whole are still too steeped in ideas from their forebears in industry, and if they are going to be successful in transforming educational systems, they need to more consciously attend to the political and relational dimensions of systemic change. Doing so would make these

methods more human; more attentive to issues of race, gender, and power; and more responsive to the rhythm and demands of public school systems, all of which are critical if these methods are going to contribute to a more just and equitable educational future. We did, however, find some examples where methods had more consciously attended to these dimensions, which we explicate in more detail below.

We develop this argument in six parts. First, we define the scope and method of our inquiry. We describe the steps we took to identify the CI methods that align with the definition above and the literature that sheds light on these methods in theory and practice, as well as some emergent patterns relating to the methods' differing theories of action and intellectual origins. This review illuminated how most CI methods are intended as responses to the complexity of education. In the second section, we elaborate on three kinds of complexity to which these methods explicitly attend: ambiguity, variability, and interdependence. Synthesizing across the intellectual forebears of CI methods, we discuss some of the perennial challenges of responding to these three types of complexity in the third section, particularly the challenges of balancing and interweaving local and expert knowledge. In Sections 4 and 5, we introduce two additional kinds of complexity stemming from the uniquely relational and political character of educational systems that are less attended to by the forebears and progenitors of CI methods. We highlight the challenges that relational and political dynamics pose for CI methods in practice and discuss the (often implicit) strategies used to grapple with these challenges. In the sixth section, we discuss how these forms of complexity pose unique challenges when considering the longer-term aims of CI methods to spur systemic change, and offer some recommendations and pathways forward. We conclude by discussing the implications of our analysis for the broader educational field.

MAPPING THE LANDSCAPE: CI METHODS AND THEIR INTELLECTUAL FOREBEARS

In approaching this chapter, we engaged in a systematic review of the literature on CI methods. We first drew on the ERIC database to search for research on CI methods that met the criteria in the introduction: methods that were grounded in local problems of practice, practitioner-centered, iterative, and focused on systems change. Through this process, we identified 14 methods that meet our criteria, and 110 empirical and theoretical publications over the past 25 years about these methods published in peer-reviewed journals and academic presses (see Supplemental Appendix A available in the online journal for additional information).¹ Table 1 provides a brief description of each method.

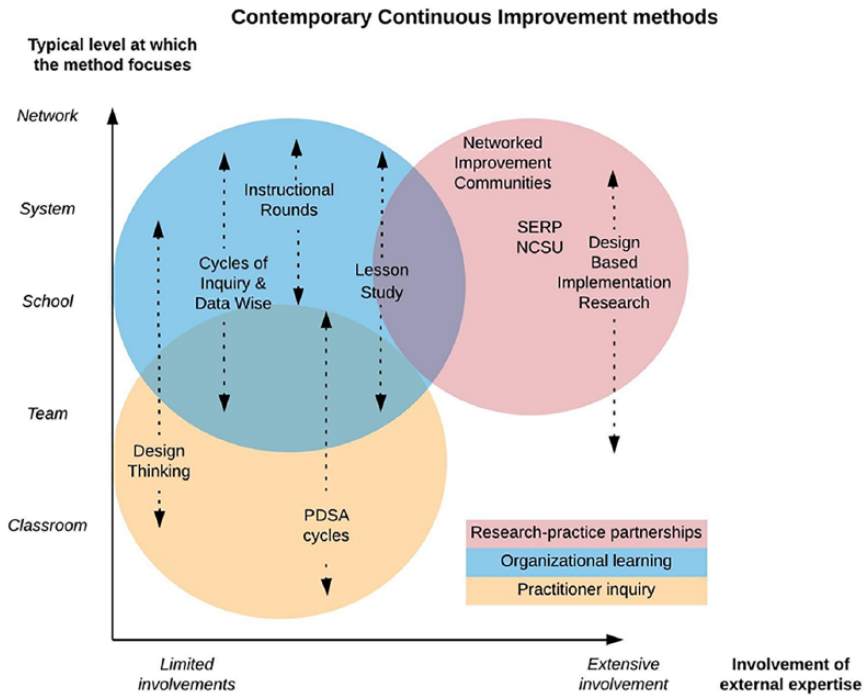
We used the theoretical articles to understand the similarities and differences across these methods' theories of action, and then drew on the empirical articles to understand the successes and challenges these methods faced in realizing their theories of action. Our reading of the literature also allowed us to organize CI methods into some rough groupings (see Figure 1). While the discussion of these methods

TABLE 1
A Brief Description of the 14 CI Methods That Served as the Focus for Our Analysis

CI Method	Target User	Description
Cycles of inquiry	Teachers	A focused, short-term, iterative process of defining a question, analyzing evidence, determining an inquiry focus, implementing an intervention, and evaluating impact
Data Teams	Teachers	A collaborative model for implementing data-driven decision making at the instructional level
The Data Wise improvement process	Teachers, teacher teams, often with school leader involvement	An eight-step process where teams of educators build a foundation for collaborative work, inquire into problems of learning and practice, and develop, implement, and continuously assess and adjust an action plan
Design-based implementation research	Combination of teachers, school leaders, researchers, and district leaders	An approach to connecting research and practice that is grounded in collaborative, iterative, systematic inquiry to develop effective, scalable, and sustainable policies and programs in schools and school districts
Design-based school improvement	School leaders	A systematic, disciplined, and design-based form of problem solving in educational organizations that involves developing and refining theories of action
Design thinking	Various	A set of mind-sets and a multistep process for redesigning a product, practice, or service around the specific needs of a user
Instructional rounds	District leaders and, school-based teams	An approach to improving instruction that involves a group of educators meeting together on an ongoing basis to define a problem of practice, visit classrooms, discuss what they saw, and identify next levels of work
Lesson study	Teachers	Originating in Japan, lesson study is a process in which teacher teams research, plan, publicly enact, discuss, and then refine intended improvements to instruction
The National Center on Scaling Up Effective Schools	Combination of teachers, school leaders, and district leaders	A continuous improvement model where researchers and practitioners co-design innovations based on outlier schools, adaptively integrate that innovation using plan-do-study-act cycles in a subset of schools, and scale up the innovation across the district
Networked Improvement Communities	Combination of teachers, school leaders, district leaders, and often outside researchers or facilitators	A network that works across sites and draws on the principles of improvement science to address a shared problem of practice using system-mapping tools and plan-do-study-act cycles
The Strategic Educational Research Partnership Spirals of inquiry	Researchers, designers, teachers, and school leaders	A research-practice partnership between researchers and designers to develop new tools or routines through iterative cycles of practitioner feedback
Teacher research	Teachers	A fluid six-phase inquiry process directing practitioners toward understanding how their students experience learning and how their own practice affects that experience
Total quality management	Combination of teachers, school leaders, and district leaders	Action-based research that enables educators to investigate questions related to teaching and learning by systematically collecting and analyzing data, reflecting on changes to their assumptions and practices, and sharing their learning
		A management approach that involves employees working on cross-department teams to continuously improve internal processes in order to enhance customer satisfaction

Note. CI = continuous improvement.

FIGURE 1
 A Venn Diagram of Some of the Continuous Improvement Methods in Our Sample, Based on the Typical Level at Which the Method Focuses and the Extent of the Involvement of External Expertise



Note. PDSA = plan-do-study-act; SERP = Strategic Educational Research Partnership; NCSU = National Center on Scaling Up Effective Schools. The placement of methods is intended to illustrate some potential differences but cannot be conclusive due to the range of examples within any given method. Note that methods can be nested inside each other, for example, PDSA cycles are part of Networked Improvement Communities.

sometimes gets subsumed under the “RPP” frame, we thought that the methods fell into three broad buckets: RPPs, organizational improvement, and practitioner inquiry. We array these as a Venn diagram to illustrate areas of overlap.²

We organized Figure 1 around two dimensions of difference that play important roles in the theories of action for each method: (a) the typical level at which the method focuses (e.g., a teacher applying the method to their own classroom vs. a whole system using the method to address a more systemic problem) and (b) the involvement of external expertise (typically researchers).

We were also interested in understanding the intellectual origins of these methods. To do this, we selected up to four articles for those methods that developed an explicit, research-based argument for their process. We identified and reviewed the major bodies of knowledge on which these methods drew (see Supplemental Appendix B available in the online journal for details on this process), keeping track of which methods cited one another and each body of knowledge. We used this analysis to identify central influences (hereafter called “forebears”) across many methods (e.g., improvement science, organizational learning), and then drew on these bodies of work to inform our analysis of these methods’ theories of action.

To create a visual representation of the themes we identified when engaging in this process, we used UCINET to conduct a social network analysis of the citation patterns between each method and these different bodies of work (Borgatti et al., 2002), which includes the extent to which these methods cited one another. We present the results of this analysis in Figure 2.

From the many connections here, we highlight a few notable themes. First, the forebears of CI methods come from a variety of sectors, disciplines, and epistemological traditions. For example, improvement science and the quality movement originated in the manufacturing sector (Deming, 1982) and then spread to health care (Cohen-Vogel et al., 2015), sociocultural theory emerged from study of apprenticeships (Lave & Wenger, 1991), and Argyris et al.’s (1985) and Argyris and Schön’s (1997) work in organizational learning and action science spanned across sectors. Interestingly, although these forebears draw heavily on the work of John Dewey—hence his centrality in our network diagram—CI methods rarely cite Dewey’s work directly. In addition, writing on “the complexity of teaching” (which represents work by scholars like David Cohen and Magdalene Lampert) was as an unexpectedly central influence across a variety of methods.

This prompted us to focus in on a couple of questions as we engaged in our review of the research:

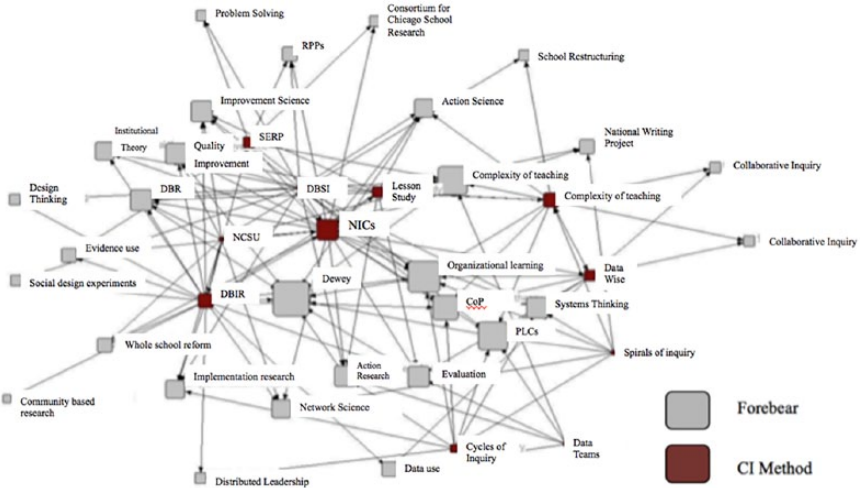
- How do the different origins of CI methods inform their theories of action and their particular affordances in the context of American education?
- What characteristics of schooling do CI methods convey when using the term *complexity*, and how have CI methods evolved from their roots in other sectors to grapple with this complexity?

We turn now to this set of questions.

CONCEPTUALIZING AND MANAGING COMPLEXITY

Our review of CI methods’ theories of action supports the finding that most methods share a conception of schooling as “complex.” This conception underpins their turn away from a “What Works” approach to education research and toward approaches that ground improvement efforts in local problems or needs, put practitioners and their perspectives at the center of research and improvement, and include

FIGURE 2
Social Network Analysis of the Citation Patterns Between and Among CI
Methods and With Intellectual Forebears



Note. CI = continuous improvement; CoP = communities of practice; DBR = design-based research; DBSI = design-based school improvement; DBIR = design-based implementation research; NCSU = National Center on Scaling Up Effective Schools; NIC = Networked Improvement Community; PLCs = professional learning communities; RPP = research–practice partnership; SERP = Strategic Educational Research Partnership. More cited forebears or methods are larger in size.

cycles of iteration. We found, however, that as we looked beyond rhetorical references to complexity to more detailed descriptions, method developers described complexity in different ways. Across the various methods we have identified three types of complexity that are inherent in the problems they seek to address. They argue that these problems are the following:

1. *Ambiguous and wicked:* Method progenitors frequently invoke complexity as it relates to ambiguity—more specifically that educational problems are often “wicked” in the sense of being ill-defined and involving competing goals or value systems (Churchman, 1967). Many progenitors ground their analysis in the complexity of teachers’ daily work, including Lampert (1985), Cohen (2011), and Dewey’s (1929/2011) earlier conception of the multiplicitous nature of teaching. Mintrop and Zumpe (2019) build on this argument, discussing how design-based approaches are particularly useful in education where “complexity and uncertainty produce ill-defined problems with unknown solution paths, unclear constraints, and ambiguous goals” (p. 304). Design methods, they argue, offer processes for “defining and framing the problem itself and considering

multiple solution paths” (p. 304). Likewise, Retna (2016) studying teachers adopting design thinking, describes design as being developed to address “ill-structured problems” (p. 6).

2. *Variable and context-specific:* A second source of complexity that progenitors seek to address is the variability within and across education systems. Variability stems from the diverse and changing needs and assets of teachers and students across grades, schools, and districts (Bryk et al., 2015; Cohen, 2011), as well as the unique local environment in which schools and classrooms are situated. For example, Penuel et al. (2011, p. 331) ground their justification for design-based implementation research (DBIR) in the varying effects of treatments across settings due to these differing local contexts.
3. *Interdependent and nested:* A third conception of complexity focuses on how the various elements of educational systems are interconnected. As a result, efforts to solve any one problem can quickly implicate many other aspects of school systems. For example, citing Rowan (2002), Penuel et al. (2011, p. 331) argue that improving educational systems “demands alignment and coordination of the actions of people, teams, and organizational units within a complex institutional ecology.” One implication of this form of complexity is that each member of the system has an inherently limited view of the problem they are trying to solve (Bryk et al., 2011). Interdependence also implies that outcomes are the product of multiple factors that interact, creating nonlinear and often unpredictable patterns of outcomes. Some methods attempt to grapple directly with this form of complexity, viewing systems as mappable and manipulable combinations of interdependent elements, such as those found in engineering (Dolle et al., 2013).

CI methods have developed many similar approaches in response to these forms of complexity. For example, they respond to variability by focusing on local problems, and address interdependence and ambiguity by gradually uncovering and provisionally addressing the inevitable unexpected outcomes within a particular context.

But these shared approaches also create numerous challenges. First, they require a delicate interweaving of formal and local knowledge—an aspiration that forebears have grappled with over the past century. In addition, they underemphasize two other aspects of complexity—which we label as the *undiscussability* of change that goes “below the green line” (Wheatley & Dalmau, 1983) and *the political fragmentation* of school systems—resulting in challenges to implementing these methods in practice. Finally, CI methods’ attention to complexity can be in tension with their ambition for scaled impact. We develop these themes in more detail below.

FORMAL AND LOCAL KNOWLEDGE: HOW SHOULD THEY RELATE?

CI methods all recognize the need to move away from an epistemology that privileges the knowledge that comes from formal, generalizable research, working instead to weave in local knowledge that can apply more closely to the specific variabilities, interdependencies, and ambiguities of particular educational

contexts. Nevertheless, they offer a range of contrasting approaches to interweaving formal and local knowledge, building on a century-long discussion about science, epistemology, and change.

Contrasting Approaches to External Knowledge

At one end of a continuum are those who are quite skeptical of the role of external knowledge. Design thinking, as a methodology, for example, celebrates the ingenuity of its practitioners, and suggests that through close observations, empathy interviews, brainstorming, and prototyping, fresh new ideas will be generated that meet the needs of the context and the moment. From this perspective external “expertise” can actually be a constraining force, as it represents the authority of the established and inhibits the creation of new insights (Kelley, 2002, pp. 64–65).

Likewise, work that emphasizes the importance of local practitioner reflection and iteration is more skeptical about formal external knowledge. Donald Schön, whose book *The Reflective Practitioner* has been highly influential in the field, argued that in the training of professionals the technical rationality of the university was overemphasized and that local craft knowledge was underemphasized (Schön, 2017). He argued that the epistemology that dominated the university, grounded in technical rationality and “application of research-based theory and technique,” was ill-suited to the kinds of problems that were most relevant to practitioners. University professors maintained their status by focusing on elegant problems and formal techniques, but real problems existed in a “swampy lowland” of ambiguity that was known to practitioners but largely invisible to most university faculty. Thus, he argued that students should apprentice under master teachers who had both extensive craft and formal knowledge, and for practitioners to engage in iterative cycles of learning and reflection as they worked.

Others have been skeptical of formal university knowledge because of its connection to histories of racism, oppression, and use to reify the educational status quo. Cochran-Smith and Lytle (1999) urge teachers to engage in ongoing critique of, and revision to, dominant ideas and resources—not only on grounds of usefulness but also on grounds of ethics, justice, ideology, and values. Much of the work on teacher action research and participatory action research (Duncan-Andrade & Morrell, 2008) takes this stance, arguing that this kind of research requires a Freirian stance to break from dominant ideologies and work with students and communities to develop alternative frames and cycles of inquiry.

More toward the middle of the spectrum are methods like Data Wise and Instructional Rounds. They do not take an explicit stance on the role of formal versus local knowledge but in practice assume that there is much local knowledge that could be mobilized to address a problem, if only a rigorous process were used to examine assumptions and generate new possibilities (Boudett et al., 2013; City et al., 2009). In so doing, they are elevating the importance of organizational knowledge *about* the process of improvement—what Deming called profound knowledge (1994)—and assuming that appropriate improvement processes will bring in both local and external knowledge as needed.

Others articulate a model that interweaves formal knowledge with local, contextual knowledge. John Dewey, for example, offered an integrative view, arguing that both were essential in supporting the variable and complex work of teaching. Dewey was opposed to simplistic ideas of knowledge transfer; anticipating much of the modern debate, he argued that local conditions were simply too complex and variable for scientific findings to be simply “applied.” He writes in *The Sources of a Science of Education* (Dewey, 1929/2011),

No conclusion of scientific research can be converted into an immediate rule of educational art. For there is no educational practice whatever which is not highly complex; that is to say, which does not contain many other conditions and factors than are included in the scientific finding. (p. 9)

Thus, he concludes, there is no substitute for the judgment and wisdom of the teacher in the moment.

At the same time, Dewey (1929/2011) argues, this judgment can be informed by the development of formal knowledge. He gives the example of a study suggesting that girls mature earlier than boys during puberty. This fact does not predetermine practice, but it helps the thoughtful teacher conceptualize his work:

The teacher who really knows this fact will have his personal attitude changed. He will be on the alert to make certain observations which would otherwise escape him; he . . . will be enabled to interpret some facts which would otherwise be confused and misunderstood. This knowledge and understanding render his practice more intelligent, more flexible and better adapted to deal effectively with concrete phenomena of practice. (p. 9)

Dewey (1929/2011) concludes therefore that while some mistakenly see science as creating uniformity: “The opposite is the case. Command of scientific methods and systematized subject-matter liberates individuals; it enables them to see new problems, devise new procedures, and, in general, makes for diversification rather than for set uniformity” (p. 6).

DBIR similarly foregrounds the interweaving of formal and local knowledge. DBIR is frequently grounded in disciplinary communities of university experts, who have spent many years thinking about how those disciplines are structured and what methods would enable teachers to teach in ambitious ways (Fishman et al., 2013). Although these experts begin with a fairly well-defined set of priors about what good instruction would look like (Penuel et al., 2011), they use codesign and adaptation processes to develop interventions that integrate their formal knowledge with the local knowledge embedded in the contexts they support (e.g., Kwon et al., 2014; Penuel et al., 2007; Severance et al., 2016).

Networked improvement communities (NICs) are perhaps the most synthetic method, offering a role for formal expertise but drawing on empathy interviews and local knowledge that comes with design thinking, and using their network structures to identify and develop positive deviants that can offer workable answers to local problems (Bryk et al., 2015). Compared to DBIR, NICs are also more flexible; they

apply profound knowledge (through the use of tools like driver diagrams, plan-do-study-act cycles [PDSAs], and root cause analysis) and then draw on different bodies of knowledge as the definition of the problem requires (Russell et al., 2017).

Which Approach Under Which Circumstances?

Three dimensions seem particularly important in thinking about which of the approaches are best suited to the problem at hand: (a) whether the local knowledge available is sufficient to tackle the problem, (b) whether the problem is clearly bounded within a single domain of existing knowledge or moves across content areas, and (c) whether the broader system authorities have an important role to play in addressing the problem.

Specifically, early research on these methods suggests that different processes are appropriate depending on whether or not local knowledge is sufficient to tackle the problem. Methods that foreground local expertise and profound knowledge can help teachers to identify problems, but those teachers can be stymied in learning about new solutions in the absence of external expertise, resulting in more superficial changes to instruction (Allen & Calhoun, 1998; Bocala, 2015; Copland, 2003; Gallimore et al., 2009; Lockwood, 2017). Methods like DBIR and lesson study that provide this external expertise but also allow for iterative adaptation have a stronger track record of addressing problems related to classroom instruction (e.g., Connor et al., 2017; Lewis & Perry, 2017).

A second dimension is flexibility. Methods like design thinking, Instructional Rounds, Data Wise, and cycles of inquiry are flexible in that they do not preassume any particular problem or any preset content expertise. Design thinking—intentionally developed in the field of design to be deployed in a great variety of different domains (Buchanan, 1992)—may be a good choice for schools and districts that are trying to innovate in more fundamental ways. From our own research, we have the example of Cowichan Valley, a district in British Columbia, which is seeking to question virtually every aspect of conventional schooling—should there be desks and chairs? what should the schedule look like?—and have found that using the design process offers a collaborative way for teachers, students, parents, and community members to develop new ideas on a variety of fronts. Methods like DBIR that foreground disciplinary content expertise are less appropriate for districts that are seeking open-ended change across a variety of dimensions.

A third dimension is the extent to which other elements of the system are implicated in the problem CI methods are seeking to address. Methods like teacher action research, Data Wise, and design thinking can be conducted locally without buy-in, coordination with, or support from top-level administrators in the system. These more local methods can energize teachers around change; can be adapted to meet the needs, interests, and goals of individual teachers; and can produce high commitment and thoughtful reflection upon practice (Copland, 2003; Goodnough, 2010; Schildkamp et al., 2016; Scribner et al., 1999; Zeichner, 2003). However, in the

absence of administrator support these efforts may be limited to coalitions of the willing, and may run into conflicts with broader policies and demands, limiting the ability to create systemic change (Artiles, 2015; Goodnough, 2010).

Conversely, larger scale RPPs are better suited to addressing more systemic problems that have garnered significant administrator buy-in. For example, NICs—with their emphasis on building system maps and driver diagrams, as well as developing diverse teams to ensure that “sufficient interest, influence, and expertise exist to address the problem” (Dolle et al., 2013, p. 447)—are well set up to support changes that require coordination across different actors in the system. However, defining problems at the system level also risks alienating teachers who have a different understanding of their most pressing needs. A fairly common challenge among RPPs that bring formal expertise to bear on a district- or network-wide problem of practice is that there are typically some practitioners who are less bought in to the framing of the problem or the pathway forward (e.g., Hannan et al., 2015; Penuel et al., 2007; Redding et al., 2018; Tichnor-Wagner et al., 2017).

In sum, like their forebears, CI methods interweave different combinations of formal and local knowledge as a way of managing the complexities of educational change. Also like their forebears, the available research on CI methods does not suggest a single “right” way of balancing or integrating formal and local knowledge. Instead, we found that different approaches may be more and less appropriate depending on the characteristics of the problem at hand. As a result, we recommend more careful consideration of and research into the fit between CI methods and the specific complexities of the problem they are trying to solve.

THE RELATIONAL ELEMENTS OF SYSTEMIC CHANGE: GOING “BELOW THE GREEN LINE” AND SURFACING THE UNDISCUSSABLE

Most CI methods acknowledge that educational improvement, relative to other domains, is a human-centric endeavor (Cohen, 1988). Indeed, the core “technology” of schooling is the ongoing interaction among teachers and students in the service of human improvement. And yet, there is little conceptualization of how CI methods might support improvement in light of the relational character of education. This absence seems to be a key consequence of the fact that many CI methods stem from fields outside of education where this relational dimension is less core to the work of improvement.

Our review suggests that to be effective, improvement efforts need to attend not simply to data, evidence, and iterative cycles but also to the relational elements of schools, which can serve as invisible enablers and barriers to change. By relational elements we refer more specifically to intersubjective understandings of teaching and learning, as well as practitioners’ individual beliefs, mind-sets, and identities, which can be replicated or disrupted through their everyday work-related actions and interactions (Smets & Jarzabkowski, 2013). In the studies of communities of practice, Wenger (1998) describes these relational elements in terms of “negotiation of

meaning”: the process by which the nature and explanation of things is defined among individuals. Relational elements also include the nature of the relationships and levels of trust among and across the different actors in the system (Bryk & Schneider, 2002). Wheatley and colleagues use the metaphor of the “green line” to distinguish between elements of organizations residing “above the green line” that are typically the focus of interventions—such as structures, operations and strategy—and the more invisible aspects of organizations (e.g., identities and relationships) that fall “below the green line” but are vital to organizational change (Wheatley, 2006; Wheatley & Dalmau, 1983).

The Necessity and Challenge of Attending to the Relational Elements of Change

As a result of a number of historical and structural features of American educational systems, the relational elements of schools can often serve as barriers to change. In particular, school systems are (a) riven with mistrust of those above in the hierarchy, especially in recent years as a result of climates of accountability and teacher evaluation (Mehta, 2013), which engenders reluctance to engage in new change efforts (Payne, 2007); (b) saddled with assumptions about teaching, professional collaboration, and racial equity that conflict with many aspirations of current reform movements (Lortie, 1975; Pollock, 2009); and (c) organized as street-level bureaucracies, wherein the impossibility of monitoring or prescribing every aspect of teachers’ work results in *de facto* autonomy for teachers, whose substantive participation in new change efforts is therefore difficult to ensure (Weatherly & Lipsky, 1977).

Unfortunately, CI methods are often not well positioned to attend to the relational elements of schools. We illustrate the tensions with examples from two educational issues that are top priorities for many using CI methods: racial equity and student-centered instruction.

First, CI methods prioritize practitioner generation and framing of problems, but issues that are in conflict with stated goals, values, and identities may not be surfaced (Argyris & Schön, 1997). Educators who believe in supporting equitable schools can still carry implicit biases that affect their practices, and teachers who aspire to improve their pedagogy may in practice have trouble giving up the belief that external factors (e.g., parental and neighborhood influence)—as opposed to their own actions as teachers—are the primary determinants of students’ achievement (Timperley & Robinson, 2001; Warikoo et al., 2016). As a result of this misalignment, the problems educators surface may not reflect the most important areas for improvement.

This challenge is compounded by the difficulty of directly observing and measuring the relational elements of schools, which can undermine CI methods’ focus on analyzable problems (Argyris et al., 1985). It is difficult to collect data on how teachers’ ongoing relationships and encounters with students can contribute to racial achievement gaps. This ambiguity allows for practitioners to suggest alternative causes of racial disparities in achievement, resulting in avoidance or confusion over

the topic (Pollock, 2009; Safir, 2017). Similarly, the differences between more and less effective approaches to student-centered learning are not easily detected through traditional assessments or brief classroom visitations (National Research Council, 2001; Spillane & Jennings, 1997).

These two challenges, combined with the value-laden nature of educators' work, make it difficult to discuss relational elements of schools openly and transparently. This poses challenges given the emphasis in CI on collaborative inquiry and decision making. For example, educators may adhere to a color-blind ideology that cautions against explicit discussions of race and the ways in which one's own biases and actions might contribute to disparate outcomes (Bonilla-Silva, 2017; Pollock, 2009). Though less fraught, it can also be difficult for teachers to surface in public discussion the ways their own and their colleagues' actions may cut against their learner-centered goals (Horn & Little, 2010; Rait, 1995). In sum, the relational elements of schools that fall "below the green line" are less visible, analyzable, and discussable, thus posing challenges for CI methods.

Going Below the Green Line: Warm and Cool Approaches

When examining the intellectual influences of CI methods, we found two strands of thinking that target "below the green line" elements of change. The first, which we characterize as the *warmer*, or *socializing*, approach, is grounded in the idea that people naturally strive to improve their craft, and the only reason they do not is because of the structures and culture of their organization. For example, Deming's (1982) theory of quality improvement tasks management with harnessing workers' natural inclination to improve by giving them opportunities to create, learn, and adjust. Theories of situated learning also reflect this approach, proposing that mindset and behavior shifts involve a gradual and natural socialization of individuals into a community with shared identity, norms, and practices (Lave & Wenger, 1991).

The second, *cooler*, or *problematizing*, approach is grounded in the idea that humans are generally resistant to change and require uncomfortable interventions to spur changes in behavior (Mezirow, 1991). Argyris and Schön (1997) find that most people in organizations prioritize winning, achieving goals, appearing rational, and minimizing negative feelings, which collectively leads to defensive behaviors and limited risk taking. They argue that real learning requires changing practitioners' assumptions, which necessitates targeting their "theories in use": making them visible, noting how they depart from their espoused theories of action, and using this disequilibrium to prompt change. Teaching on adaptive leadership often takes a similar approach (e.g., Heifetz et al., 2009).

In theory, these two approaches can and should be integrated. Building trust and shared norms creates the space and psychological safety that enables people to have hard conversations about problematic assumptions, norms, and practices (Edmondson & Lei, 2014; Safir, 2017). However, in practice, we found that CI methods rarely attended explicitly to both warmer and cooler approaches to change. A number of methods, particularly those structured as RPPs, prioritize building shared norms and

trust between researchers and practitioners and creating a climate where educators feel safe to take risks in trying out new ideas (Coburn & Penuel, 2016; Donovan et al., 2013; Penuel, 2015). However, these methods rarely utilize that foundation to push educators to reevaluate existing beliefs about their work. For example, Roegman et al. (2017) described how the efforts of university-based facilitators to lead inquiry into more sensitive topics were stymied by a prevailing “culture of nice” that limited discussions of systemic inequities.

In contrast, the PDSA process embedded in NICs and the National Center on Scaling Up Effective Schools, as well as many data use protocols, helps make visible and challenge educators’ assumptions but includes fewer explicit strategies for building shared norms and culture. For example, educators engaged in PDSA cycles often assimilate the cycles into their existing, less well developed, approach to improvement, which can prevent them from using the process to destabilize and address other assumptions about their practice (Hannan et al., 2015; Tichnor-Wagner et al., 2017). PDSA cycles are a form of counternormative work (Hannan et al., 2015), and educators are not used to seeing their own successes and challenges as important knowledge that could guide others’ work, if well documented (Tichnor-Wagner et al., 2017). One study of design-based school improvement similarly found that educators struggled to use the process in a way that did not just confirm their existing ideas about improvement, despite sustained support from researchers around this challenge (Mintrop & Zumpe, 2019). And in a review of data use in schools, Datnow and Park (2018) found that efforts to use data for CI often focused instead on using it only to group students or meet narrow accountability demands.

There are some exceptions to this either/or pattern. Traditions of teacher research (e.g., Cochran-Smith & Lytle, 1990) support teacher teams in taking on fundamentally different practices, norms, and identities that relate to their roles as knowledge producers in order to support more critical inquiry into their own practice and prevailing educational ideas and values. Data Wise transitions from a socializing approach early on to a problematizing approach later in the cycle (Boudett et al., 2013). For example, Data Wise offers activities for educators to better understand who they are in relation to their team (e.g., the “compass points” protocol) so that in later steps, with the help of protocols designed to minimize risk or vulnerability for individual teachers, teams are better able to have difficult conversations about instructional problems of practice. But little research has examined the success of these methods in garnering shifts in teachers’ thinking and practice (for exceptions, see Allen & Calhoun, 1998; Goodnough, 2010).

Overall, CI methods could benefit from attending more to the relational elements of schooling that fall “below the green line” but are necessary for deep and sustained change (Safir, 2017). This might best be accomplished by intentionally fostering both socializing (warmer) and problematizing (cooler) approaches to change, striking a balance between safety and challenge (Edmondson, 2002). This could involve drawing more explicitly on forebears that integrate these approaches to change,

including traditions of teacher research, and adapting methods that were designed in other sectors to better meet the relational demands of educational contexts.

POLITICAL FRAGMENTATION: HIGH DEMAND, TURBULENCE, AND INCOHERENCE

A second source of complexity overlooked by CI progenitors is political fragmentation. The American political context of local democratic control over education produces layers of educational governance that offer competing and often conflicting imperatives (Cohen & Spillane, 1992; Chubb & Moe, 1990). Perhaps because CI methods stem in part from manufacturing, on the one side, or classroom-based design, on the other, the methods have to date provided less guidance in how to manage the political complexities that are an inevitable feature of change efforts in a democratic public school system. The particular challenges with which CI methods must grapple can be summarized as high demand, turbulence, and incoherence.

High demand stems from the need for educators to be responsive not only to the needs of their students and families but also to the expectations stemming from other layers of the fragmented educational system (Chubb & Moe, 1990). As a result, the already demanding work of teaching is often compounded by a torrent of time-consuming external expectations (e.g., frequent interim and state assessments, evaluation procedures, school improvement planning, mandatory professional development), all in the context of limited preparation and collaborative time (Johnson, 2013; Kraft et al., 2015). In this time- and energy-scarce environment, CI methods ask teachers and leaders to take on new roles and responsibilities that, however potentially rich and meaningful, may risk as being seen as another distraction if not carefully integrated into their existing work. This in turn may contribute to burnout (Martin & Gobstein, 2015) or more superficial engagement with CI methods, as practitioners prioritize those demands that are familiar or come with sanctions (Leary et al., 2016).

Turbulence enters into school systems, particularly those in underresourced communities, through the high rates of teacher, principal, and superintendent churn, and with it waves of partially implemented reforms (Hess, 2011; Payne, 2008). Such changes cut against the stability needed for CI methods to flourish (Glazer & Peurach, 2013), and may contribute to reform fatigue and ceremonial compliance, while also impeding efforts to build a culture supportive of CI methods (Coburn et al., 2013; Englert et al., 1977; Rosenquist et al., 2015). CI methods may unintentionally exacerbate this turbulence because funding for this work is often temporary, in which case it risks becoming yet another briefly implemented reform (Detert et al., 2000; Leary et al., 2016; Martin & Gobstein, 2015).

A final challenge is *incoherence*. As a result of the contested purposes of education and the varied stakeholders whom schools are meant to serve, educators often experience an incoherent environment that pulls them in many competing directions at once (Cohen & Spillane, 1992; Ingersoll, 2005; Lampert, 1985). These competing values and priorities can push against improvement efforts, such as when districts and states implement CI methods in ways that conform to the piecemeal, rushed, and

compliance-oriented approach of educational bureaucracies that methods are meant to subvert (Akiba & Wilkinson, 2016).

CI method progenitors have not explicitly attended to how their methods address these challenges of political fragmentation. However, our review identified three different, more implicit ways these methods try to manage fragmentation, each with certain weaknesses.

Approach 1: Head Down, Ignoring the System: Equipping Educators With Inquiry Tools

This approach attempts to isolate improvement efforts from broader political dynamics, equipping educators with inquiry tools needed to understand and improve the core work of teaching and learning. Methods like Data Wise, Data Teams, and lesson study seek to build certain organizational preconditions for successful inquiry work but then direct the bulk of their attention toward teaching practice and student understanding, with a focus on problems of practice that are within a teacher's immediate control (Boudett et al., 2013).

This approach is vulnerable to political fragmentation. At the school level, educators' take-up of inquiry processes is frequently constrained by a multiplicity of other initiatives (Tichnor-Wagner et al., 2017), lack of ongoing access to relevant data (Hannan et al., 2015), and districts' history with past reforms (Scribner et al., 1999), which can create a climate of distrust that makes inquiry and learning more difficult (Allen & Calhoun, 1998; Ingram et al., 2004). For example, despite a competitive application process, coaches in Russell et al.'s (2019) study struggled to document their PDSA cycles due to the number of other expectations they had to manage. Copland (2003) found that teachers facing external pressure for improvement who had limited time to engage in inquiry would often jump to solutions before understanding the problems they were trying to solve. As a result of these challenges, many studies have pointed to the important role of school-based leadership and advocacy in supporting this work (e.g., Copland, 2003; Hannan et al., 2015; Perry & Lewis, 2009), which in turn can be cultivated by having district leaders directly engaging in these methods as well (Cannata et al., 2017; Lockwood et al., 2017; Rigby et al., 2018; Roegman et al., 2015).

Approach 2: Designing Coherent Niches or Subsystems

A second approach to managing fragmentation involves the creation of niches (Cohen & Mehta, 2017) or subsystems that are both coherent and adaptable to teachers' local context. This approach is typically used among methods in the DBIR tradition (Frumin, 2019). For example, Anderson et al. (2018) described their aim as "creating 'tool kits' for a curricular activity system" that includes aligned teacher and students guides, professional development and assessments. This approach takes on a fairly narrow slice of the work of a school system such as a biology curriculum or earth science unit and then develops the kind of coherent

and adaptable infrastructure that school systems typically lack around that “slice.” Lesson study takes a similar approach, frequently coupling together the lesson study cycle with the kinds of conceptually rich curriculum materials that are present in Japan but lacking in the United States (Lewis & Perry, 2017).

To manage the political environment, such niches frequently have to engage in *buffering*—protecting their work from the demands of the larger systems—and *bridging*—finding ways to connect their work to the demands of the larger system. For example, when an RPP in Denver Public Schools began piloting a biology curriculum aligned to the Next Generation Science Standards, teachers initially received lower ratings on their formal evaluations due to evaluators not being familiar with the new standards. The RPP developed a new observation protocol but had to continuously engage in this kind of bridging and buffering in order to create coherence for teachers (Frumin, 2019; Honig & Hatch, 2004; Penuel, 2015).

Scholars have also stressed that these kinds of subsystems require a stable and supportive system environment—a precondition less likely to be available in low-performing districts (Debarger et al., 2013). Researchers in the design-based tradition emphasize the importance of engaging in CI to better negotiate these challenges (e.g., Anderson et al., 2018), but the methods and process for this kind of second-order CI are less fleshed out (for an exception, see Peurach et al., 2016).

Approach 3: Systematic Attention to the Environment

The third, and most comprehensive, approach is to work directly with actors in the larger education system. For example, in a NIC, network members are asked to map out the entire system contributing to the problem of practice and potential drivers of improvement. By bringing together a variety of critical stakeholders, NICs also aim to bridge communication and coordination problems that arise as a result of fragmentation. The NCSU model is similar in many ways to NICs, except that rather than mapping out the full system, they focus on identifying distinguishing practices from positive outlier schools and using those to inform the district-level and school-level plans (Cohen-Vogel et al., 2016). When there is buy-in across the system, there are clear advantages to this approach in terms of comprehensiveness and reducing some of the conflicting imperatives described above.

But this approach also comes with trade-offs. Given the tacit and relational dimensions of change, building a shared driver diagram or coordinated approach to improvement does not necessarily ensure that teachers and those actually implementing the reforms are fully on board with the change. Indeed, a persistent difficulty across NICs and NCSUs is for educators to connect their school-level improvement work with the broader aims of the network (Cannata et al; 2017; Martin & Gobstein, 2015; Tichnor-Wagner et al., 2017). Understanding that these maps will be used in politicized contexts with multiple, competing interest groups, we also wonder if they privilege what is legible, politically safe, and representative of dominant interests. For example, in our study of an early-literacy NIC, we found

participants turned away from a focus on the impact of trauma on students' engagement with literacy toward data-driven guided reading groups. The latter approach enabled the NIC to focus on a well-bounded problem of practice that was well aligned to the political aims of the county department of education, but it also resulted in a lost opportunity to fully address literacy challenges stemming specifically from students' experience with trauma.

In sum, CI methods operate not in closed systems but instead in systems that must be responsive to a dynamic and complex ecosystem of policies, reforms, and intermediary organizations (Burch, 2007; Rowan, 2002). We worry that without explicitly attending to this environment, CI methods risk contributing to the very incoherence they are trying to ameliorate. This risk is particularly great as CI receives increased attention from funders and policymakers, which increases the temptation to use these methods in ritualized or ceremonial ways that may please external stakeholders while exacerbating demands on educators (Yurkofsky, 2017; Peurach et al., 2018). Avoiding this outcome will likely require more explicit attention to how CI methods operate given fragmentation. For example, we have wondered whether it is preferable to intentionally embed CI methods within district or state governance as a way of minimizing incoherence and multiple demands, or if that risks corruption of the CI process by politics and accountability (e.g., Datnow & Park, 2018).

IMPACT AMID COMPLEXITY: CHALLENGES AND PATHWAYS FORWARD

Challenges

CI methods face an immense challenge in their efforts to address educational problems that manifest the five forms of complexity—ambiguity, variability, interdependence, indiscussability, and political fragmentation—that we have laid out in this chapter. When looking across CI methods' theories of action, we identified four different visions for influencing the educational sector despite this complexity. As above, these different visions are implied rather than explicitly justified:

- One vision for large-scale impact involves embedding collaborative inquiry processes into the work of teaching and leadership. We see this approach in most localized CI methods. For example, Lewis et al. (2006) articulate how lesson study is *not* about developing increasingly refined interventions but rather about deepening educators' knowledge, commitment, and learning resources through ongoing communal inquiry.
- A second vision involves school systems drawing on improvement science to enhance reform efforts. This approach is best seen in NICs and the NCSU, and involves educators using PDSA cycles to modify new tools or initiatives so that they are better integrated into educators' work (Hannan et al., 2015), and

network leaders using these insights to improve the reform effort (Redding et al., 2018).

- A third vision involves using collaborative processes like codesign to develop interventions that are better tailored to the needs of teachers (Penuel et al., 2007). This approach involves designing interventions (and complementary systems of adult learning) that support educators in making productive adaptations (Scherrer et al., 2013). This allows for interventions to be scaled up while still being useful to teachers with varied needs and in diverse contexts.
- A fourth vision, which often runs alongside the others described, involves the spread of RPPs, such that practitioners have better access to research as they embark on improvements, and researchers can generate better theory through sustained and iterative collaboration.

These methods thus imply substantial changes in the roles of teachers, leaders, and researchers; the way districts organize for improvement; the process by which instructional materials are designed; and the relationship between school systems and universities. Even though most of these methods are still in their first or second decade, there is already some promising evidence that each of these approaches to large-scale impact could be successful. Evidence from a randomized control trial of lesson study across 39 educator teams suggests that collaborative inquiry processes can improve teaching and learning when coupled with instructional resources (Lewis & Perry, 2017). Yamada et al. (2018) used propensity score matching to demonstrate the positive impact of their NIC on improving completion rates of remedial math courses for more than 4,000 students across 10 community colleges. In addition, a number of interventions that have come out of DBIR (all of which are RPPs) have also led to improved teaching practice and student outcomes (Connor, et al., 2017; DeBarger et al., 2017; Donovan & Snow, 2018; Wright & Gotwals, 2017).

Inattention to these five forms of complexity, however, may frustrate CI methods' efforts at larger scale impact. Efforts to scale up collaborative inquiry have faced challenges related to education's ambiguous causal relationships and flawed outcome measurements (Copland, 2003; Ingram et al., 2004), sensitive and value-laden topics of inquiry (Roegman et al., 2017), and conflicting external demands (Hubers et al., 2017). In practice, this can result in educators departing from CI's core principles, such as in making decisions via intuition over evidence, not assessing the effectiveness of chosen interventions, or acting and identifying solutions before understanding the problem (Allen & Calhoun, 1998; Copland, 2003; Detert et al., 2000; Hubers et al., 2017; Mintrop & Zumpe, 2019; Nelson, 2009). Efforts to use improvement science to support the scaling up of reforms have similarly been constrained in practice by political fragmentation (Hannan et al., 2015; Russell et al., 2019; Tichnor-Wagner et al., 2017). For example, Redding and Viano (2018), studying a networked partnership, found that in the hopes of garnering staff buy-in and addressing the prior

histories of reform efforts, teacher leaders developed less disruptive changes that were also less likely to substantially improve teaching.

Efforts to collaboratively design more educator-centered interventions also struggled to confront these different sources of complexity. Anderson et al. (2018) offer one of the few accounts of an attempt to codesign tools that would scale up to the level of “thousands of classrooms” (p. 1028), finding that implementation varied across classrooms and that the prior classroom culture prevailed in determining whether tools were implemented as intended. In the Strategic Education Research Partnership with the Boston Public Schools, researchers found that only “internally coherent” schools could integrate their initially designed literacy tools. This led to a refocusing around the design of additional tools to measure and improve internal coherence (Donovan & Snow, 2018).

Pathways Forward

Across the different visions for scaling impact, CI methods often struggle to integrate into the complex systems they are seeking to transform. We see four necessary steps to take this field forward.

The first tasks researchers with clarifying and testing out the different mechanisms by which CI methods yield improvements in teaching and learning (Lochmiller & Lester, 2017). Underlying each CI method is an organizational theory of change (however implicit or incomplete) that is grounded in assumptions about how people learn, work together, and solve problems of varying complexity. Drawing on Weiss’s (1995) work on theory-driven evaluation, we believe much can be learned by making these theories explicit and then evaluating them (e.g., Perry & Lewis, 2009), focusing not just on the overall effect of CI methods on student outcomes but also on the effects of certain components of a CI method’s theory of action on more intermediate outcomes (e.g., related to internal coherence, professional community, or social network structure). Although there are not yet many areas of sustained and comparative investigation into the key mechanisms by which CI methods improve teaching and learning, we reviewed a number of stand-alone studies that serve as great starting points for such research. For example, in studying CI in a school district, Redding and Viano (2018) tested the theory that teacher ownership of a scale-up process improves implementation. Likewise, Hatch et al. (2016) draw on social network analysis to evaluate the mechanisms by which instructional rounds might support the development of a community of practice.

Second, we recommend investigating CI methods in terms of the combination of protocols and routines that make them up. This would involve fewer research questions about the impact of a given CI method, and more questions about which components of methods are most appropriate for a given context or problem. We believe more openness toward *cross-pollination* could be helpful. By cross-pollination, we mean thinking of CI methods as bundles of different routines, processes, and

strategies, each with a particular purpose that could be appropriate depending on the context or problem. For example, PDSAs are a prominent protocol for iteration across CI methods, but educators have struggled to use PDSAs in certain contexts, particularly when problems are less well defined and data sources more problematic. We therefore would encourage exploring other methods of iteration that might be better tailored to more ambiguous educational contexts, such as the research lesson process of lesson study. A cross-pollination approach would also help to leverage the complementary strengths and weaknesses of different CI methods that we have identified in this chapter.

Cross-pollination might also allow for greater flexibility in the role of researchers and a questioning of RPPs as a defining feature of many CI methods. Although there are certainly contexts in which the prominent role of researchers seems essential (e.g., the codesigning of new curriculum and assessments), researcher involvement brings unique challenges (Coburn et al., 2013). Depending on the problem at hand, these challenges may outweigh the expertise researchers offer.

Clarifying the role researchers should play in CI methods across different contexts and over time relates to our third recommendation, which is to specify—even as tentative hypotheses—the underlying theory of action for how CI methods will yield impact in the long-term. Although many CI methods are funded through generous and short-term grants, we found little discussion of how improvement efforts might persist or expand when these additional funds run out.³ Moreover, empirical research into CI methods tends to focus on whether and how these methods work under ideal conditions, rather than how they might scale into new contexts or with less funding. For example, while there is quasi-experimental evidence that the Carnegie Math Pathways—an intervention that includes courses, professional development, student support, and participation in an NIC—increased completion of developmental math requirements (Yamada et al., 2018), we now wonder how this intervention could be scaled up without losing integrity. How essential is it that the program continues to exist as part of a network, or to draw on improvement science? More broadly, while much research has documented the early stages of RPP work (Cobb et al., 2013; DeBarger et al., 2017; Penuel et al., 2007; Wright & Gotwals, 2017), less is known about how well designers can foresee and design for the variable contexts in which educators may use these materials as it spreads out and/or scales up over time.

A final recommendation is for leaders in the CI tradition to recognize their dual roles as institutional entrepreneurs working to build a new field of research and improvement that runs counter to many of the structures, norms, and assumptions of the status quo (Peurach et al., 2018), and researchers engaging in rigorous work of advancing knowledge and improving outcomes within this new field. These two roles are essential, but they often push in competing directions. Our worry is that too much of a field-building orientation might divert attention

from identifying and improving upon the limitations of these methods, which we believe is necessary for the long-term success of this movement (Peterson, 2016).

CONCLUSION

Continuous improvement, in all its guises, is becoming increasingly influential in education reform, including state and federal policy, district and school improvement plans, as well as in the language of many foundations and educational nonprofits. While the roots of it are old, and connect to industry and design as well as education, its logic has been newly embraced in recent years. But while in one sense what it offers is anodyne—who could be against “continuous improvement”?—our excavation suggests that its underlying ideas are actually quite radical in their intentions and aspirations. The purveyors of such movements are seeking to move away from top-down policy, to help teachers and school leaders embrace evidence and work more scientifically, to change the relationship between researchers and practitioners, and to surface and confront deep underlying issues of inequity. These more radical goals are at odds with many of the structural and cultural features of the American educational system, and thus despite the unexceptionable name, what we are really witnessing is a deeply countercultural movement that challenges, and seeks to transform, many aspects of this system.

We foresee two possible scenarios. The first is that the language of CI gets assimilated into existing ways of doing things. Districts adopt “continuous improvement” but do so within a paradigm of compliance and control, teachers adopt inquiry cycles but inquire in ways that are consistent with their pedagogical priors, and researchers and schools work together in ways that allow them to win grants and produce publications but do not lead to deep improvements in practice. This pattern of assimilation would be consistent with previous countercultural efforts at instructional reform (Cohen, 1990), and it is the most likely outcome. We fear that what would happen under this scenario is that data and inquiry cycles would become the myth and ceremony of the modern age, widely institutionalized but leading to little change either in practice or in the fundamental routines, structures, processes, and culture that govern the sector (Yurkofsky, 2017).

A second scenario is more hopeful. Under this view, CI processes recognize that what they are promoting is counter to dominant logics but gradually and practically create new ways of working that are consistent with the deeper aspirations of the movement. As people experience these new routines, identities, and roles, they experience greater success and efficacy than they did in the past, and, in so doing, their appetite for doing more work under this new paradigm increases. That creates the motivation for more work of this new type, and over time, the work deepens as people become more familiar with the new way of working and become increasingly knowledgeable and skilled.


For this second scenario to come to pass, this review suggests that movement proponents will need to become more forthright than they have been so far in taking on

the challenges that the educational sector presents. In particular, they need to work actively to manage the political turbulence, incoherence, and conflicting demands that are so characteristic of the American education sector, and they will need to create humane CI processes that mix warm and cool approaches, and go “below the green line” to address many of the unseen, adaptive elements that are critical to deep and sustainable change. In so doing, they could help CI processes become the disruptive force that is inherent in their deepest aspirations, as opposed to becoming yet another reform that is swallowed by the forces it is seeking to transform.

ACKNOWLEDGMENTS

We would like to thank the reviewers and editors for their invaluable feedback throughout the revision process. The ideas in this chapter were originally presented at the 2019 Meeting of the American Educational Research Association, during which time we received helpful feedback from our discussant, Christopher Redding, as well as many others who attended. We are also grateful to David Sherer who provided feedback on an earlier draft of this chapter. Work on this chapter was supported by a research grant from the Spencer Foundation (No. 256554).

ORCID iD

Amelia Peterson  <https://orcid.org/0000-0002-7225-9068>

NOTES

¹We limited our sample to published research, excluding conference papers. While conference papers can be useful at illustrating ideas that have not yet made it into published research, for this review we were interested in taking stock of the field and its evolution over time, and thus limited our scope to work that has been published.

²This map displays a subset of the CI methods in our sample in order to illustrate our argument. This map is not intended to be comprehensive.

³These observations were informed by a discussion at the 2019 American Educational Research Association session “Building Knowledge About Research-Practice Partnerships” where we presented an early draft of these arguments. Christopher Redding and Joshua Glazer, in particular, provided insight on this point.

REFERENCES

- Akiba, M., & Wilkinson, B. (2016). Adopting an international innovation for teacher professional development: State and district approaches to lesson study in Florida. *Journal of Teacher Education, 67*(1), 74–93. <https://doi.org/10.1177/0022487115593603>
- Allen, L., & Calhoun, E. F. (1998). Schoolwide action research: Findings from six years of study. *Phi Delta Kappan, 79*(9), 706–710. <https://www.jstor.org/stable/20439316>
- Anderson, C. W., de los Santos, E. X., Bodbyl, S., Covitt, B. A., Edwards, K. D., Hancock, J. B., Lin, Q., Thomas, C. M., Penuel, W. R., & Welch, M. M. (2018). Designing educational systems to support enactment of the Next Generation Science Standards. *Journal of Research in Science Teaching, 55*(7), 1026–1052. <https://doi.org/10.1002/tea.21484>
- Argyris, C., Putnam, R., & Smith, D. M. (1985). *Action science* (Vol. 13). Jossey-Bass.

- Argyris, C., & Schön, D. A. (1997). Organizational learning: A theory of action perspective. *Rei: Revista Española de Investigaciones Sociológicas*, (77/78), 345–348. <https://doi.org/10.2307/40183951>
- Artiles, J. A. (2015). *The Education DesignShop: Broadening non-designers' solutions for big issues* [Master's thesis, Massachusetts Institute of Technology]. <http://dspace.mit.edu/handle/1721.1/103568>
- Bocala, C. (2015). From experience to expertise: The development of teachers' learning in lesson study. *Journal of Teacher Education*, 66(4), 349–362. <https://doi.org/10.1177/0022487115592032>
- Bonilla-Silva, E. (2017). What we were, what we are, and what we should be: The racial problem of American sociology. *Social Problems*, 64(2), 179–187. <https://doi.org/10.1093/socpro/spx006>
- Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). *UCINET for Windows: Software for social network analysis* [Computer software]. Analytic Technologies.
- Boudett, K. P., City, E. A., & Murnane, R. J. (2013). *Data wise: A step-by-step guide to using assessment results to improve teaching and learning*. Harvard Education Press.
- Bryk, A. S., Gomez, L. M., & Grunow, A. (2011). Getting ideas into action: Building networked improvement communities in education. In M. T. Hallinan (Ed.), *Frontiers in sociology of education* (pp. 127–162). Springer. https://doi.org/10.1007/978-94-007-1576-9_7
- 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*. Harvard Education Press.
- Bryk, A. S., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. Russell Sage Foundation.
- Buchanan, R. (1992). Wicked problems in design thinking. *Design Issues*, 8(2), 5–21. <https://doi.org/10.2307/1511637>
- Burch, P. (2007). Educational policy and practice from the perspective of institutional theory: Crafting a wider lens. *Educational Researcher*, 36(2), 84–95. <https://doi.org/10.3102/0013189X07299792>
- Cannata, M., Cohen-Vogel, L., & Sorum, M. (2017). Partnering for improvement: Improvement communities and their role in scale up. *Peabody Journal of Education*, 92(5), 569–588. <https://doi.org/10.1080/0161956X.2017.1368633>
- Chubb, J. E., & Moe, T. M. (1990). *Politics, markets, and America's schools*. Brookings Institution Press.
- Churchman, C. (1967). Wicked problems [Guest editorial]. *Management Science*, 14(4), 141–146. <https://doi.org/10.1287/mnsc.14.4.B141>
- City, E. A., Elmore, R. F., Fiarman, S. E., & Teitel, L. (2009). *Instructional rounds in education*. Harvard Education Press.
- 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. *National Society for the Study of Education Yearbook*, 112(2), 320–349.
- 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). *Practice partnerships: A strategy for leveraging research for education*. W. T. Grant Foundation.
- Cochran-Smith, M., & Lytle, S. L. (1999). Chapter 8: Relationships of knowledge and practice: Teacher learning in communities. *Review of Research in Education*, 24(1), 249–305. <https://doi.org/10.3102/0091732X024001249>
- Cohen, D. K. (1988). Teaching practice: Plus ça change. In P. W. Jackson (Ed.), *Contributing to educational change: Perspectives on research and practice* (pp. 27–84). McCutchan.

- Cohen, D. (1990). A revolution in one classroom: The case of Mrs Oublier. *Education Evaluation and Policy Analysis, 12*(3), 311–329.
- Cohen, D. K. (2011). *Teaching and its predicaments*. Harvard University Press. <https://doi.org/10.4159/harvard.9780674062788>
- Cohen, D. K., & Mehta, J. D. (2017). Why reform sometimes succeeds: Understanding the conditions that produce reforms that last. *American Educational Research Journal, 54*(4), 644–690. <https://doi.org/10.3102/0002831217700078>
- Cohen, D. K., & Spillane, J. P. (1992). Policy and practice: The relations between governance and instruction. *Review of Research in Education, 18*, 3–50. <https://doi.org/10.2307/1167296>
- Cohen-Vogel, L., Cannata, M., Rutledge, S. A., & Socol, A. R. (2016). A model of continuous improvement in high schools: A process for research, innovation design, implementation, and scale. *Teachers College Record, 118*(13).
- 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, 29*(1), 257–277. <https://doi.org/10.1177/0895904814560886>
- Connor, C. M., Dombek, J., Crowe, E. C., Spencer, M., Tighe, E. L., Coffinger, S., Zargar, E., Wood, T., & Petscher, Y. (2017). Acquiring science and social studies knowledge in kindergarten through fourth grade: Conceptualization, design, implementation, and efficacy testing of content-area literacy instruction (CALI). *Journal of educational psychology, 109*(3), 301–320. <https://doi.org/10.1037/edu0000128>
- Copland, M. A. (2003). Leadership of inquiry: Building and sustaining capacity for school improvement. *Educational Evaluation and Policy Analysis, 25*(4), 375–395. <https://doi.org/10.3102/01623737025004375>
- Datnow, A., & Park, V. (2018). Opening or closing doors for students? Equity and data use in schools. *Journal of Educational Change, 19*(2), 131–152. <https://doi.org/10.1007/s10833-018-9323-6>
- Debarger, A. H., Choppin, J., Beauvineau, Y., & Moorthy, S. (2013). Designing for productive adaptations of curriculum interventions. *National Society for the Study of Education Yearbook, 112*(2), 298–319.
- Debarger, A. H., Penuel, W. R., Moorthy, S., Beauvineau, Y., Kennedy, C. A., & Boscardin, C. K. (2017). Investigating purposeful science curriculum adaptation as a strategy to improve teaching and learning. *Science Education, 101*(1), 66–98. <https://doi.org/10.1002/scs.21249>
- Deming, W. E. (1982). *Out of the crisis*. MIT press.
- Deming, W. E. (1994). *The new economics for industry, government, education*. MIT press.
- Detert, J. R., Kopel, M. B., Mauriel, J., & Jenni, R. (2000). Quality management in U.S. high schools: Evidence from the field. *Journal of School Leadership, 10*(2), 158–87. <https://doi.org/10.1177/105268460001000203>
- Dewey, J. (2011). *The sources of a science of education*. Martino Fine Books. (Original work published 1929)
- 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 Yearbook, 112*(2), 443–463.
- Donovan, M. S., & Snow, C. (2018). Sustaining research–practice partnerships: Benefits and challenges of a long-term research and development agenda. In B. Bevan & W. R. Penuel (Eds.), *Connecting research and practice for educational improvement* (pp. 33–50). Routledge. <https://doi.org/10.4324/9781315268309-3>
- Donovan, M. S., Snow, C., & Daro, P. (2013). The SERP approach to problem-solving research, development, and implementation. *National Society for the Study of Education Yearbook, 112*(2), 400–425.

- Duncan-Andrade, J. M. R., & Morrell, E. (2008). *The art of critical pedagogy: Possibilities for moving from theory to practice in urban schools* (Vol. 285). Peter Lang. <https://doi.org/10.3726/b12771>
- Edmondson, A. C. (2002). The local and variegated nature of learning in organizations: A group-level perspective. *Organization Science*, 13(2), 128–146. <https://doi.org/10.1287/orsc.13.2.128.530>
- Edmondson, A. C., & Lei, Z. (2014). Psychological safety: The history, renaissance, and future of an interpersonal construct. *Annual Review of Organizational Psychology and Organizational Behavior*, 1(1), 23–43. <https://doi.org/10.1146/annurev-orgpsych-031413-091305>
- Englert, R. M., Kean, M. H., & Scribner, J. D. (1977). Politics of program evaluation in large city school districts. *Education and Urban Society*, 9(4), 429–450. <https://doi.org/10.1177/001312457700900403>
- 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. *National Society for the Study of Education*, 112(2), 136–156.
- Frumin, K. (2019). Researchers and practitioners in partnership: Co-design of a high school biology curriculum [Unpublished doctoral dissertation]. Harvard Graduate School of Education.
- Gallimore, R., Ermeling, B. A., Saunders, W. M., & Goldenberg, C. (2009). Moving the learning of teaching closer to practice: Teacher education implications of school-based inquiry teams. *The Elementary School Journal*, 109(5), 537–553. <https://doi.org/10.1086/597001>
- Glazer, J. L., & Peurach, D. J. (2013). School improvement networks as a strategy for large-scale education reform: The role of educational environments. *Educational Policy*, 27(4), 676–710. <https://doi.org/10.1177/0895904811429283>
- Goodnough, K. (2010). Teacher learning and collaborative action research: Generating a “knowledge-of-practice” in the context of science education. *Journal of Science Teacher Education*, 21(8), 917–935. <https://doi.org/10.1007/s10972-010-9215-y>
- 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., Hill, K., & Roegman, R. (2016). Investigating the role of instructional rounds in the development of social networks and district-wide improvement. *American Educational Research Journal*, 53(4), 1022–1053. <https://doi.org/10.3102/0002831216653205>
- Heifetz, R. A., Grashow, A., & Linsky, M. (2009). *The practice of adaptive leadership: Tools and tactics for changing your organization and the world*. Brighton, MA: Harvard Business Press.
- Hess, F. M. (2011). *Spinning wheels: The politics of urban school reform*. Washington, DC: Brookings Institution Press.
- Honig, M. I., & Hatch, T. C. (2004). Crafting coherence: How schools strategically manage multiple, external demands. *Educational Researcher*, 33(8), 16–30. <https://doi.org/10.3102/0013189X033008016>
- Horn, I. S., & Little, J. W. (2010). Attending to problems of practice: Routines and resources for professional learning in teachers’ workplace interactions. *American Educational Research Journal*, 47(1), 181–217. <https://doi.org/10.3102/0002831209345158>
- Hubers, M. D., Schildkamp, K., Poortman, C. L., & Pieters, J. M. (2017). The quest for sustained data use: Developing organizational routines. *Teaching and Teacher Education*, 67, 509–521. <https://doi.org/10.1016/j.tate.2017.07.007>
- Ingram, D., Louis, K. S., & Schroeder, R. G. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *Teachers College Record*, 106(6), 1258–1287. <https://doi.org/10.1111/j.1467-9620.2004.00379.x>

- Johnson, C. C. (2013). Educational turbulence: The influence of macro and micro-policy on science education reform. *Journal of Science Teacher Education, 24*(4), 693–715. <https://doi.org/10.1007/s10972-012-9333-9>
- Kelley, T. (2002). *The art of innovation: Lessons in creativity from IDEO, America's leading design firm: Success through innovation the IDEO way* (Main ed.). Profile Books.
- Kraft, M. A., Papay, J. P., Johnson, S. M., Charner-Laird, M., Ng, M., & Reinhorn, S. (2015). Educating amid uncertainty: The organizational supports teachers need to serve students in high-poverty, urban schools. *Educational Administration Quarterly, 51*(5), 753–790. <https://doi.org/10.1177/0013161X15607617>
- Kwon, S. M., Wardrip, P. S., & Gomez, L. M. (2014). Co-design of interdisciplinary projects as a mechanism for school capacity growth. *Improving Schools, 17*(1), 54–71. <https://doi.org/10.1177/1365480213519517>
- Lampert, M. (1985). How do teachers manage to teach? Perspectives on problems in practice. *Harvard Educational Review, 55*(2), 178–195. <https://doi.org/10.17763/haer.55.2.56142234616x4352>
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511815355>
- Leary, H., Severance, S., Penuel, W. R., Quigley, D., Sumner, T., & Devaul, H. (2016). Designing a deeply digital science curriculum: Supporting teacher learning and implementation with organizing technologies. *Journal of Science Teacher Education, 27*(1), 61–77. <https://doi.org/10.1007/s10972-016-9452-9>
- Lewis, C., & Perry, R. (2017). Lesson study to scale up research-based knowledge: A randomized, controlled trial of fractions learning. *Journal for Research in Mathematics Education, 48*(3), 261–299. <https://doi.org/10.5951/jresmetheduc.48.3.0261>
- Lewis, C., Perry, R., & Murata, A. (2006). How should research contribute to instructional improvement? The case of lesson study. *Educational Researcher, 35*(3), 3–14. <https://doi.org/10.3102/0013189X035003003>
- Lochmiller, C. R., & Lester, J. N. (2017). Conceptualizing practitioner-scholarship for educational leadership research and practice. *Journal of Research on Leadership Education, 12*(1), 3–25. <https://doi.org/10.1177/1942775116668525>
- Lockwood, M., Dillman, M., & Boudett, K. P. (2017). Using data wisely at the system level. *Phi Delta Kappan, 99*(1), 25–30. <https://doi.org/10.1177/0031721717728275>
- Lockwood, M. G. (2017). *Refining the art of coaching: Organizational learning on a district data inquiry team* [Unpublished doctoral dissertation]. Harvard Graduate School of Education.
- Lortie, D. C. (1975). *Schoolteacher: A sociological study* (Vol. 21). University of Chicago Press.
- Martin, W. G., & Gobstein, H. (2015). Generating a networked improvement community to improve secondary mathematics teacher preparation: Network leadership, organization, and operation. *Journal of Teacher Education, 66*(5), 482–493. <https://doi.org/10.1177/0022487115602312>
- Mehta, J. (2013). *The allure of order: High hopes, dashed expectations, and the troubled quest to remake American schooling*. Oxford University Press.
- Mezirow, J. (1991). *Transformative dimensions of adult learning*. Jossey Bass.
- Mintrop, R., & Zumpe, E. (2019). Solving real-life problems of practice and education leaders' school improvement mind-set. *American Journal of Education, 125*(3), 295–344. <https://doi.org/10.1086/702733>
- National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. National Academies Press.
- Nelson, T. H. (2009). Teachers' collaborative inquiry and professional growth: Should we be optimistic? *Science Education, 93*(3), 548–580. <https://doi.org/10.1002/sc.20302>

- Payne, C. M. (2008). *So much reform, so little change: The persistence of failure in urban schools*. Cambridge, MA: Harvard Education Press.
- Payne, C. M. (2007). *I've got the light of freedom: The organizing tradition and the Mississippi freedom struggle*. University of California Press.
- Penuel, W. R. (2015, September 22–25). *Infrastructuring as a practice for promoting transformation and equity in design-based implementation research* [Keynote presentation]. International Society for Design and Development in Education Conference, Boulder, CO, United States.
- Penuel, W. R., Fishman, B. J., Haugan Cheng, B., & 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>
- Penuel, W. R., Peurach, D. J., LeBoeuf, W. A., Riedy, R., Barber, M., Clark, T., & Gabriele, K. (2018). *Defining collaborative problem solving research: Common values and distinctive approaches*. LearnDBIR.
- Penuel, W. R., Roschelle, J., & Shechtman, N. (2007). Designing formative assessment software with teachers: An analysis of the co-design process. *Research and Practice in Technology Enhanced Learning*, 2(1), 51–74. <https://doi.org/10.1142/S1793206807000300>
- Perry, R. R., & Lewis, C. C. (2009). What is successful adaptation of lesson study in the US? *Journal of Educational Change*, 10(4), 365–391. <https://doi.org/10.1007/s10833-008-9069-7>
- Peterson, A. (2016). Getting ‘What Works’ working: Building blocks for the integration of experimental and improvement science. *International Journal of Research & Method in Education*, 39(3), 299–313. <https://doi.org/10.1080/1743727X.2016.1170114>
- Peurach, D. J., Lenhoff, S. W., & Glazer, J. L. (2016). Large-scale high school reform through school improvement networks: Exploring possibilities for “developmental evaluation.” *Teachers College Record*, 118(13).
- Peurach, D. J., Penuel, W. R., & Russell, J. L. (2018). Beyond ritualized rationality: Organizational dynamics of instructionally-focused continuous improvement. In M. Connolly, D. H. Eddy-Spicer, C. James, & S. D. Hruse (Eds.), *The Sage handbook of school organization* (pp. 465–488). Sage. <https://doi.org/10.4135/9781526465542.n28>
- Pollock, M. (2009). *Colormute: Race talk dilemmas in an American school*. Princeton University Press. <https://doi.org/10.1515/9781400826124>
- Rait, E. (1995). Against the current: Organizational learning in schools. In S. B. Bacharach & B. Mundell (Eds.), *Images of schools: Structures and roles in organizational behavior* (pp. 71–107). Corwin.
- Redding, C., & Viano, S. L. (2018). Co-creating school innovations: Should self-determination be a component of school improvement? *Teachers College Record*, 120(11).
- Redding, C., Cannata, M., & Miller, J. (2018). System learning in an urban school district: A case study of intra-district learning. *Journal of Educational Change*, 19(1), 77–101. <https://doi.org/10.1007/s10833-017-9310-3>
- Retna, K. S. (2016). Thinking about “design thinking”: A study of teacher experiences. *Asia Pacific Journal of Education*, 36(Suppl. 1), 5–19. <https://doi.org/10.1080/02188791.2015.1005049>
- Rigby, J. G., Forman, S., Fox, A., & Kazemi, E. (2018). Leadership development through design and experimentation: Learning in a research–practice partnership. *Journal of Research on Leadership Education*, 13(3), 316–339. <https://doi.org/10.1177/1942775118776009>
- Roegman, R., Hatch, T., Hill, K., & Kniewel, V. S. (2015). Relationships, instruction, understandings: One district’s implementation of rounds. *Journal of Educational Administration*, 53(5), 625–641. <https://doi.org/10.1108/JEA-07-2014-0078>

- Roegman, R., Allen, D., & Hatch, T. (2017). The elusiveness of equity: Evolution of instructional rounds in a superintendents network. *American Journal of Education, 124*(1), 127–159. <https://doi.org/10.1086/693957>
- Rosenquist, B. A., Henrick, E. C., & Smith, T. M. (2015). Research–practice partnerships to support the development of high quality mathematics instruction for all students. *Journal of Education for Students Placed at Risk, 20*(1–2), 42–57. <https://doi.org/10.1080/10824669.2014.988335>
- Rowan, B. (2002). The ecology of school improvement: Notes on the school improvement industry in the United States. *Journal of Educational Change, 3*(3–4), 283–314. <https://doi.org/10.1023/A:1021277712833>
- 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*(5).
- Russell, J. L., Correnti, R., Stein, M. K., Bill, V., Hannan, M., Schwartz, N., Booker, L. N., Pratt, N. R., & Matthis, C. (2019). Learning from adaptation to support instructional improvement at scale: Understanding coach adaptation in the TN Mathematics Coaching Project. *American Educational Research Journal*. <https://doi.org/10.3102/0002831219854050>
- Safir, S. (2017). *The listening leader: Creating the conditions for equitable school transformation*. John Wiley.
- Scherrer, J., Israel, N., & Resnick, L. B. (2013). Beyond classrooms: Scaling and sustaining instructional innovations. *Yearbook of the National Society for the Study of Education, 112*(2), 426–442.
- Schildkamp, K., Poortman, C. L., & Handelzalts, A. (2016). Data teams for school improvement. *School Effectiveness and School Improvement, 27*(2), 228–254. <https://doi.org/10.1080/09243453.2015.1056192>
- Schön, D. A. (2017). *The reflective practitioner: How professionals think in action*. Routledge. <https://doi.org/10.4324/9781315237473>
- Scribner, J. P., Cockrell, K. S., Cockrell, D. H., & Valentine, J. W. (1999). Creating professional communities in schools through organizational learning: An evaluation of a school improvement process. *Educational Administration Quarterly, 35*(1), 130–160. <https://doi.org/10.1177/0013161X99351007>
- Severance, S., Penuel, W. R., Sumner, T., & Leary, H. (2016). Organizing for teacher agency in curricular co-design. *Journal of the Learning Sciences, 25*(4), 531–564. <https://doi.org/10.1080/10508406.2016.1207541>
- Smets, M., & Jarzabkowski, P. (2013). Reconstructing institutional complexity in practice: A relational model of institutional work and complexity. *Human Relations, 66*(10), 1279–1309. <https://doi.org/10.1177/0018726712471407>
- Spillane, J. P., & Jennings, N. E. (1997). Aligned instructional policy and ambitious pedagogy: Exploring instructional reform from the classroom perspective. *Teachers College Record, 98*(3), 449–81.
- Tichnor-Wagner, A., Wachen, J., Cannata, M., & Cohen-Vogel, L. (2017). Continuous improvement in the public school context: Understanding how educators respond to plan–do–study–act cycles. *Journal of Educational Change, 18*(4), 465–494. <https://doi.org/10.1007/s10833-017-9301-4>
- Timperley, H. S., & Robinson, V. M. (2001). Achieving school improvement through challenging and changing teachers' schema. *Journal of Educational Change, 2*(4), 281–300. <https://doi.org/10.1023/A:1014646624263>
- Warikoo, N., Sinclair, S., Fei, J., & Jacoby-Senghor, D. (2016). Examining racial bias in education: A new approach. *Educational Researcher, 45*(9), 508–514. <https://doi.org/10.3102/0013189X16683408>

- Weatherly, R., & Lipsky, M. (1977). Street-level bureaucrats and institutional innovation: Implementing special-education reform. *Harvard Educational Review*, 47(2), 171–197. <https://doi.org/10.17763/haer.47.2.v870r1v16786270x>
- Weiss, C. H. (1995). Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families. In J. P. Connell, A. C. Kubisch, L. B. Schorr, & C. H. Weiss (Eds.), *New approaches to evaluating community initiatives: Concepts, methods, and contexts* (1st ed., pp. 65–92). Aspen Institute.
- Wenger, E. (1998). Communities of practice: Learning as a social system. *Systems Thinker*, 9(5), 2–3.
- Wheatley, M. J. (2006). *Leadership and the new science: Discovering order in a chaotic world* (3rd ed.). Berrett-Koehler.
- Wheatley, M. J., & Dalmau, T. (1983). *Below the green line or the 6 circle model* [Unpublished manuscript].
- Wright, T. S., & Gotwals, A. W. (2017). Supporting kindergartners' science talk in the context of an integrated science and disciplinary literacy curriculum. *The Elementary School Journal*, 117(3), 513–537. <https://doi.org/10.1086/690273>
- Yamada, H., Bohannon, A. X., Grunow, A., & Thorn, C. A. (2018). Assessing the effectiveness of Quantway®: A multilevel model with propensity score matching. *Community College Review*, 46(3), 257–287. <https://doi.org/10.1177/0091552118771754>
- Yurkofsky, M. M. (2017). *The restructuring of educational organizations: From ceremonial rules to technical ceremonies* [Unpublished doctoral dissertation]. Harvard Graduate School of Education.
- Zeichner, K. M. (2003). Teacher research as professional development for P–12 educators in the USA [1]. *Educational Action Research*, 11(2), 301–326. <https://doi.org/10.1080/09650790300200211>