Continuous Improvement in Education Settings

A Literature Review

Michael S. Garet, Sara Mitrano, Ryan Eisner, Julie R. Kochanek, Kathleen T. Jones, Maggi Ibis, and Scott Estrada

AUGUST 2021

American Institutes for Research® | AIR.ORG

AIR®
Advancing Evidence.
Improving Lives.

Continuous Improvement in Education Settings

A Literature Review

AUGUST 2021

Authors:
Michael S. Garet, Sara Mitrano, Ryan Eisner, Julie R. Kochanek, Kathleen T. Jones, Maggi Ibis, and Scott Estrada

American Institutes for Research®
1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
202.403.5000 | **AIR.ORG**

This report is based on research funded by the Bill & Melinda Gates Foundation. The findings and conclusions contained within are those of the authors and do not necessarily reflect positions or policies of the Bill & Melinda Gates Foundation.

15793_03/22

Contents

Exhibits

Acknowledgments

The authors wish to thank Elizabeth Nolan and Liz Scalia who provided initial guidance and structure for this review. We also thank Larry Friedman, PhD, whose review of this literature review improved its content. Finally, we thank the Publication and Creative Services staff at AIR who edited and formatted the review in its final stages.

Executive Summary

Over the last decade, educators have become increasingly interested in continuous improvement (CI) as a strategy for reform. CI requires practitioners to engage in iterative cycles of inquiry by defining local problems of practice, testing potential interventions, studying the results, and improving upon those interventions. This method of improvement stands in contrast to approaches focused on adopting "off the shelf" research-based practices or evidence-based interventions.

Although there is a large literature outlining the rationale for CI and specifying methods and tools that support its use, the empirical literature on CI in education is far smaller and less easily accessible. Thus, we set out to locate empirical studies that described the process of implementing CI, as well as studies that analyzed the conditions that foster or limit its use. We identified 34 empirical studies that met our final eligibility criteria (detailed in Appendix A).

Key findings:

- Almost all of the reviewed sources were case studies and small in scope. Most studies reported on CI work undertaken by a single school, district, or another organizational unit for a 1- or 2-year time period.
- Although iterative inquiry cycles are a fundamental component of CI, relatively few studies provided detailed data on the implementation of cycles.
- There is some evidence that engaging in CI may be associated with improved student outcomes, but the evidence is based on weak research designs. Only three studies employed a comparison group, and no study was based on random assignment.
- The studies suggest that implementing CI is facilitated by supportive leadership, opportunities to collaborate, and the provision of professional development. But these conclusions are based largely on perceptions of the CI participants. Only one study explicitly compared conditions associated with high- and low-implementing CI teams.

Our results suggest that CI is a promising avenue for reform, but much more work is necessary to assemble evidence on the use of iterative inquiry cycles to improve student outcomes. One challenge is that this approach to CI is a local effort embedded in practice, not an approach that involves researchers testing the efficacy of standardized interventions. To build an evidence base for CI, some CI efforts must be undertaken within a systematic research framework that not only supports local learning by the CI team, but also supports drawing generalizable conclusions about CI by researchers, practitioners, and policymakers.

Introduction

Over the last decade, practitioners and researchers have become increasingly interested in data-driven approaches to school improvement. More specifically, there is growing interest in approaches that require practitioners to engage in iterative cycles of inquiry by defining localized problems of practice, testing potential interventions, studying the results, and improving upon those interventions. This method of improvement stands in contrast to approaches focused on adopting "off the shelf" research-based practices or evidence-based interventions.

Work in this new tradition is sometimes labeled continuous improvement (hereafter, CI), improvement science, or quality improvement (Berwick, 1989). While scholars specify the key components of CI differently, the methods that have been proposed have much in common (Cohen-Vogel et al., 2015; Park et al., 2013).[1] The key ideas are encapsulated in the Model of Improvement (Langley et al., 2009), which advises local improvement teams to begin their work by focusing on three questions:

1. What are we trying to accomplish?
2. How will we know that a change is an improvement?
3. What changes can we make that will result in an improvement?

The notion is that work should begin by identifying an aim and the measures that will be used to assess improvement, then settling on change ideas to test. Once change ideas are identified, the Model of Improvement calls for change ideas to be tested in a series of Plan-Do-Study-Act (PDSA) cycles. PDSA cycles, which originated in industry in the work of Deming (1986) and Shewhart (1991), reflect the fundamental premise that CI rests on iterative testing.[2] In each cycle, teams plan a specific change idea to test, do the work specified in the test, study the results, and then act on them.

The Bill & Melinda Gates Foundation developed a related set of ideas as part of its Networks for School Improvement (NSI) initiative (NSI, 2019). The NSI conception is grounded in six core parameters of CI:

- An understanding of the problem, the systems that produce current inequitable outcomes, and the opportunities and assets of the community and their students.

[1] For discussions of the history of CI, alternative models, and its application to education, see Cohen-Vogel et al. (2015) and Park et al. (2013).
[2] For more information, see Deming (1986) and Taylor et al. (2014).

- A clear and specific aim centered on achieving equitable outcomes for Black, Latino/a, and low-income students.
- An equity-centered theory of practice improvement for how to reach the aim.
- Disciplined inquiry cycles to test interventions and collect and analyze data to assess if changes are an improvement.
- Collaborative and diverse teams comprised of people with time, expertise, experience, and will to tackle the problem.
- Use of locally relevant and valued data from multiple sources, relevant research, and measurement as keys to improvement.

Although there is a large literature outlining the rationale for CI and specifying methods and tools that support its use, the empirical literature on CI in education is far smaller and less easily accessible. Thus, we set out to locate published empirical studies (including journal articles, dissertations, and reports) that described implementation of CI, as well as studies analyzing the conditions that foster or limit its use. In addition, we looked for studies that report whether teaching or learning outcomes changed over the course of improvement work. We limited the search to CI in educational organizations: schools, districts, and postsecondary institutions. (The empirical literature in the health care field is more substantial; see, for example, Taylor et al., 2014.)

Recent work in education has argued that CI can be strengthened by situating the work in planned social networks or collaborative research partnerships that connect practitioners with similar aims to engage in work on similar problems. Networks can support teams in sharing lessons and, in particular, enable teams to learn from variation in outcomes across settings. Bryk et al. (2015) laid out six principles for networked improvement communities, building on the Model for Improvement:

- Make the work problem-specific and user-centered.
- Variation in performance is the core problem to address.
- See the system that produces the current outcomes.
- We cannot improve at scale what we cannot measure.
- Anchor practice improvement in disciplined inquiry.
- Accelerate improvements through networked communities.

Although CI efforts in education often occur in networked settings, we focused our review on the CI aspects of the work (e.g., the implementation of PDSA cycles), which has received less

attention in the literature. Two recent reviews examined the literature on improvement networks in detail: Barletta et al. (2018) and Bush-Mecenas et al. (2020).

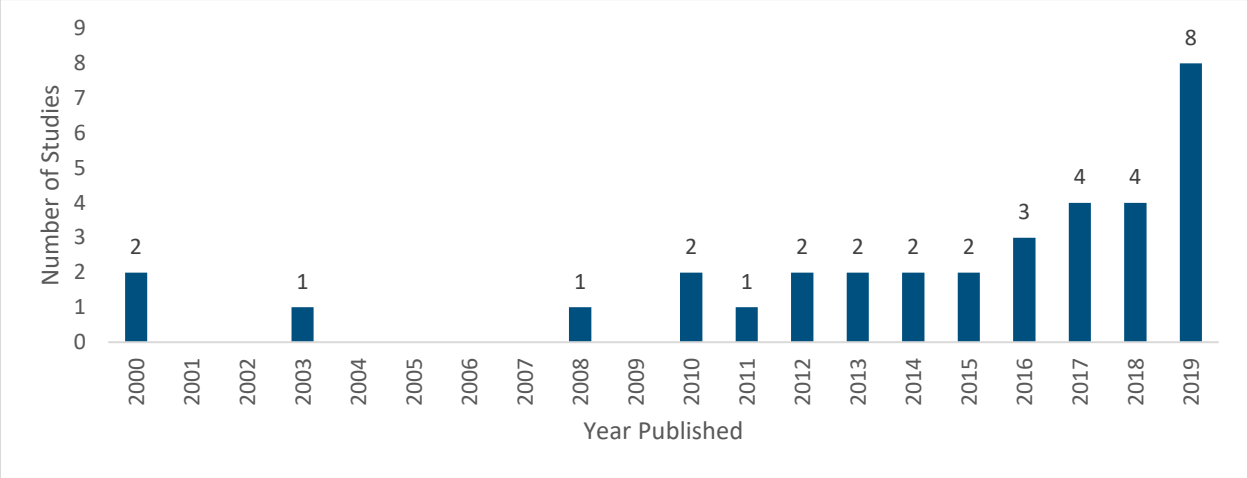
We began the review with a systematic search of CI studies published since 1995. (See Appendix A for a discussion of the search and screening methods.) We searched for empirical studies that reported on CI, such as case studies or research reports. We excluded articles that were primarily conceptual, including, for example, those that discussed the history, motivation, or rationale for CI as well as handbooks that provided materials and supports for those engaging in CI work. The conceptual literature on CI often references examples of CI in action, for example, providing information on the types of data used to identify the problem of practice or used during PDSA cycles (Bryk, 2018; Park et al., 2013) or providing case examples of organizations engaged in CI. We excluded these examples from our review unless they were based on systematic data collected on an improvement effort.

We identified 34 empirical studies that met our eligibility criteria (see Appendix B for a list of the 34 studies along with their basic features). Sixteen studies were published in journals; the others were research reports or dissertations (see Exhibit 1). Most studies were published after 2010 (see Exhibit 2).

Exhibit 1. About Half of Studies Were Nonjournal Studies

Publication type	Number of studies
Journal	16
Report	10
Dissertation	8

Exhibit 2. Most Studies Were Published After 2010



Twenty-five of the 34 studies we reviewed described distinct CI initiatives. The other nine studies described three initiatives: the CORE-PACE research partnership (five studies), the National Center on Scaling Up Effective Schools (two studies), and the Quest Project (two studies). Appendix C lists the nine studies that focused on these three initiatives.

We focused our review of each study on seven key issues:

- **The ways the studies defined CI**. In our search, we found that many studies that used the phrase "continuous improvement" did not use it in a sense consistent with the Model of Improvement; that is, they did not focus on iterative inquiry cycles. The boundary between studies of CI and other related approaches to reform (e.g., data use, professional learning communities, or the measurement of implementation fidelity) is not always clear.
- **The research designs employed**. We examined whether the studies were primarily descriptive, the scale of the studies (e.g., the number of schools or other organizations involved) and whether the studies were explicitly comparative (e.g., comparing outcomes in schools implementing CI with other schools).
- **The duration, setting, and focus of the CI initiatives**. We were particularly interested in the time covered by the studies: Did they examine one or multiple years of CI?
- **The interventions and CI process**. One core focus of our review is the implementation of inquiry cycles—in particular, the change ideas that were the focus of inquiry and the number of cycles completed.
- **The outcomes of CI**. We asked whether studies reported on the outcomes achieved during the CI process, and, if so, whether there was evidence of improvement.
- **Supportive conditions and barriers of CI**. We examined whether the studies gathered information on conditions supporting or inhibiting the implementation of CI.
- **The diffusion of CI**. Finally, we examined whether the studies reported on if practices diffused within and beyond the schools or other organizations initially involved in the CI effort.

The sections that follow discuss each of these issues.

The Ways the Studies Defined Continuous Improvement

How Did the Reviewed Studies Define "Continuous Improvement"?

Although the practice of continuous improvement (CI) is becoming increasingly popular among educators, there is no universal definition of "continuous improvement." Identifying studies that focused on CI proved challenging because authors frequently used the term in a broad sense, to connote a focus on school improvement or quality, without necessarily indicating an effort to implement inquiry cycles or other features associated with the Model of Improvement or similar approaches. For example, Pate et al. (2012) conducted in-depth case studies to examine how three organizations serving disconnected youth[3] used data for continuous program improvement. However, the process described in the study did not involve staff conducting iterative inquiry cycles; staff used existing data tools and systems to monitor youth progress. In addition, the boundaries between CI and other approaches to reform—for example, professional learning communities (Coburn & Russell, 2008), research-practice partnerships (Coburn et al., 2013), data use (Boudett et al., 2013), and implementation science (Eccles & Mittman, 2006)—are blurry.

We established two eligibility criteria for studies to be included in the review. We restricted the review to studies that

- focused on CI initiatives carried out by practitioners (e.g., teachers, school-level administrators) and
- discussed practitioners implementing one or more inquiry cycles (e.g., PDSA cycles).

The first criterion ruled out studies that described researcher efforts to improve an intervention before testing it. The second criterion ruled out studies that characterized CI as planning for school improvement or using annual accountability data to assess school progress but did not involve conducting inquiry cycles. Although we did not require studies to offer an explicit definition of CI to be included in the review, authors of 27 of the 34 studies defined CI, commonly referencing the Bryk et al. (2015) principles for networked improvement communities or the Deming (1986) PDSA cycle.

[3] In the study, Pate et al. (2012) defined the term "disconnected youth" as young adults "ages 14-24, who are not employed, not on-track to graduate high school, and/or who have dropped out of high school altogether" (p. 3).

The following examples illustrate how some study authors defined "continuous improvement":

> **Ray (2014, p. 12)**: Continuous improvement is an ongoing process used to improve services, products, or processes through a specific focus on data to drive organizational decisions. Many organizations use the plan-do-check-act (PDCA) cycle, which is also known as the Deming Cycle or Shewhart Cycle.
>
> **Harrison et al. (2019, p. 2)**: Continuous improvement approaches have gained attention by emphasizing structured, iterative processes of design, implementation, and evaluation that are responsive to individual school contexts, promote organizational learning over time, and embrace both the wisdom of practice and empirical research (Bryk et al., 2015; Cohen-Vogel et al., 2015; Lewis, 2015; Sabelli & Harris, 2015; Shulman, 1986).
>
> **Aumiller (2008, p. 13)**: For an academically high-performing district, improving test scores is not the overarching goal. Rather, implementing an effective continuous improvement model to affect systemic change is the challenge. Continuous improvement demands relentless pursuit of doing things better, more efficiently, at lower costs, and with greater sensitivity to the customers' needs. Ultimately, it will distinguish the leading educational systems of the 21st century (Kimmelman & Kroeze, 2002).
>
> **Cannata et al. (2016, p. 1)**: One such approach, improvement science, focuses on using continuous improvement processes to address problems of practice in ways that contribute to organizational learning (Bryk et al., 2015). Key to improvement science is the creation of networked improvement communities that engage in disciplined cycles of inquiry where data is used to understand the problem and test potential solutions (Bryk et al., 2011).

As another way to characterize the ways study authors defined CI, we created a word frequency query to display the most frequently used words in the authors' descriptions of CI.[4] We identified 14 terms, which we classified into six groups. Exhibit 3 lists the characteristics or features from most to least frequently mentioned.

[4] We used the Query Wizard function in NVivo, a qualitative data analysis computer software, to identify keywords within each study's text that described CI. We grouped words with the same stem (e.g., grouping together "test," "tested," and "testing"). We eliminated pronouns, numbers, and words that lacked descriptive characteristics (e.g., design).

Exhibit 3. Characteristics or Features of CI Definitions Found in Studies

Characteristic or feature	Description	Keywords (stemmed words)	Number of studies that referenced keywords[a]
Disciplined inquiry	Conduct inquiry cycles, also known as PDSA cycles, which are a repeated, iterative process of improvement	Act (acts, acting) Cycle (cycles) Inquiry Plan (plans, planning)	22
Systems focused	Focus on improving the systems that are producing current outcomes	System (systems, systemic)	18
Problem focused	Identify or discover a problem of practice based on improvement needs	Need (needs, needed) Problem (problems)	16
Innovation	Create an innovation (i.e., intervention or change idea) to solve a problem and test that innovation	Idea (ideas) Test (tests, testing, tested)	15
Data driven	Use and analyze data and evidence to identify and/or address the problem	Data Evidence	13
User centered	Engage practitioners in the process of addressing the problem and considering the local context	Context (contexts) Practitioner (practitioners)	12

[a] Some studies are classified by identifying multiple characteristics or features.

The Research Designs Employed in the Studies

One reason we undertook the review was to learn about the research designs used to study the implementation of continuous improvement (CI). Although CI entails gathering data as a component of the work, our review focused on the data collected by researchers about the implementation and effects of CI as a reform. Almost all of the studies we reviewed were case studies, providing descriptive information on the CI teams, processes, and outcomes. Only three of the 34 studies employed comparison groups. Arbour et al. (2016) introduced a new preschool curriculum in 40 centers; 18 of the 40 centers received support to engage in CI as part of the implementation process, and the results were compared in the two groups using propensity weights to adjust for preexisting differences in the centers. Brown (2013) examined a set of low-performing schools that had been offered support for CI. The author compared three schools that fully implemented the CI process, three partial implementers, and three schools that did not implement CI. Finally, Gallagher et al. (2019) compared schools that launched Local Improvement Teams (LITs) to schools that did not have LITs.

Most studies were small in scope (see Exhibit 4). Fourteen studies focused on CI work undertaken in a single school, district, or another organizational unit. Only four studies included 21 or more schools, limiting the information available on the variation across schools in the implementation and effects of CI.

Exhibit 4. Most Studies Reported Data on Five or Fewer CI Teams

Number of teams engaged in CI	Number of studies
1	14
2–5	12
6–10	3
11–20	1
21–30	2
31–40	1
41–50	1

Authors' Involvement in Continuous Improvement Work

Researchers often are involved as partners on CI teams to develop interventions or change ideas and tools to measure improvement. Researchers who are involved as partners likely have access to the data generated during CI; thus, they are in a position to provide detailed information on the CI process. However, external researchers might be viewed as more objective.

Nineteen of the 34 studies included in the review had authors who were involved in the CI work (see Exhibit 5). The authors' involvement in the CI work included facilitating CI team discussions, being an active member of the CI team, leading the network of CI teams, and/or developing tools that CI team members could use to implement an intervention. Some authors created the CI framework used as part of the CI initiative; for example, two studies (Howley-Rowe, 2000a, 2000b) were authored by AEL Inc., which is the organization that created the CI framework implemented by school teams in the Quest Project.

Fifteen of the 19 studies whose authors were involved in the CI work described a CI effort conducted through research-practice partnerships. For example, one study discussed how researchers from Virginia Tech University collaborated with staff from the Department of Defense Education Activity (DoDEA) to help DoDEA become a CI organization (Kaufman et al., 2019). Two studies described the research-practice partnership work of the National Center on Scaling Up Effective Schools, a project funded by the Institute of Education Sciences (Harrison et al., 2019; Tichnor-Wagner et al., 2017).

Exhibit 5. About Half of Studies' Authors Were Involved in the CI Work

Authors involved in the CI work	Number of studies
Yes	19
No	15

What Types of Data Did the Study Researchers Collect to Study the Implementation of Continuous Improvement?

Thirty of the 34 studies explicitly described the types of data collected by the study authors to study CI implementation. Of these 30 studies, seven studies included authors who were involved in the CI effort; for example, as a facilitator for CI teams (Arnold, 2010; Ray, 2014) or as a leader of the CI network (Rohanna, 2018). Exhibit 6 shows that practitioner interview data were the most common source of data used in the studies, followed by practitioner observations and artifact analysis. Eleven studies collected data on student academic achievement.

Exhibit 6. Practitioner Interview Data Were the Most Common Source Collected by Study Authors

Data type	Number of studies that listed data type	Example(s) of data types from studies
Practitioner interview data	26	Administrator interviews, teacher interviews, and focus groups with CI team members
Practitioner observations	21	Classroom observations to identify trends in instructional practices, CI team meeting observations, and network convening observations
Artifact analysis	18	Network convening presentation materials, agendas from CI team meetings, and PDSA tracking logs
Practitioner survey data	14	Survey to collect data related to individual and organizational improvement science capacity
Student academic data	11	Standardized test scores and benchmarking assessment tool

The Duration, Setting, and Focus of Continuous Improvement

For How Many Years Was the Continuous Improvement Effort Underway?

Two key questions might be asked about the duration of continuous improvement (CI) efforts:

- How long does it take for organizations to build the capacity to carry out CI effectively?
- Does the capacity to engage in CI improve with experience?

Half of the studies reported on a CI effort that was in its first or second year (see Exhibit 7). Most other studies reported on CI efforts lasting 2 or 3 years. None of the studies explicitly examined whether the implementation of CI deepened with years of experience.

Exhibit 7. About Half of Studies Examined CI Efforts That Took Place Over 1 or 2 Years

Duration	Number of studies
1 year	13
2 years	7
3 years	7
4 years	2
5 years	3
6 years	1
15 years	1

Where Did the Continuous Improvement Effort Take Place?

We examined several aspects of the settings described in the studies. First, we assessed the organizational level at which the effort took place (e.g., school or district). We primarily classified studies' organizational level by examining which members were part of the CI team(s); for studies that focused on Grades K–12 students, we looked at which members were school- or district-level staff. As indicated in Exhibit 8, about half of the studies described CI efforts conducted at the school level only, and slightly less than half of the studies were conducted at both the school and district levels.

Exhibit 8. The Majority of Studies Described CI Efforts Conducted at the School and/or District Levels

Organizational level(s)	Number of studies	Example from studies
School only	15	Four high school science teachers worked together to implement inquiry cycles with support from school administrators (Ermeling, 2010).
School and district	13	The CI effort was a collaboration between district- and school-level innovation design teams. Within this model, district staff served as "critical friends" for each school team (Harrison et al., 2019).
District only	1	A district team used its data system and a locally developed online tool to plan for and test improvements that were designed to increase the number of eligible high school students applying to multiple colleges and universities (Aguilar et al., 2017).
Postsecondary	4	A group of teacher educators implemented a form of lesson study to improve their curriculum and student knowledge (Liang et al., 2019).
Federal government	1	A government agency that operates one of two federally operated school systems engaged in various CI efforts (Kaufman et al., 2019).

The studies varied in the focal grade bands (see Exhibit 9). Eight studies focused on CI efforts in high school, and 17 described CI efforts conducted across multiple grade bands. For example, Brown (2013) examined data from three high schools, three middle schools, and three elementary schools.

Exhibit 9. Half of the Studies Described CI Efforts Conducted Across Multiple Grade Bands

Grade band	Number of studies
Prekindergarten only	1
Elementary school (Grades K–5) only	3
Middle school (Grades 6–8) only	1
High school (Grades 9–12) only	8
Multiple	17
Not applicable[a]	4

[a] Studies that described CI efforts conducted in postsecondary settings were categorized as "not applicable."

What Types of Students Were the Focus of the Continuous Improvement Effort?

Twenty-six of the 34 studies described the characteristics of students who were the focus of the CI effort (see Exhibit 10). Fifteen studies described CI efforts focused on serving students of color or diverse student populations. Few studies described CI efforts focused specifically on improving learning outcomes for Black (*n* = 4) or Latino/a/x students (*n* = 5).

Exhibit 10. Students of Color Were The Most Common Focal Population of CI Work in Studies

Focal population	Number of studies[a]	Example descriptors from studies
Students of color or diverse student populations	15	• African American students • Culturally, economically, and linguistically diverse • Ethnically diverse • Hispanic students • Latinx students • Students of color • Majority/minority district • Minority served
Students experiencing poverty	12	• Economically disadvantaged students • Low-income students • Priority students • Students who qualified for free or reduced-price lunch • Students in poverty
Students from low-performing schools or districts	9	• Low, lower, or lowest performing districts or schools • Schools that were “struggling” • Schools with “lower than average proficiency rates”
Students from high-performing schools or districts	4	• High-performing schools or districts • Performance “higher than overall district”
Other	3	• Students in special education • Military-connected students

[a] Some studies are classified as describing multiple types of students who were the focus of CI efforts (e.g., students experiencing poverty and students of color).

What Was the Focus of the Continuous Improvement Effort?

In two thirds of the studies, the CI effort focused on improving student achievement; in the other third, the effort focused on improving teacher practice (see Exhibit 11). Fewer studies described CI efforts focused on students' social-emotional outcomes, access to postsecondary education, or family engagement.

Exhibit 11. Improving Students' Academic Skills Was The Most Common Focus of the CI Effort in Studies

CI effort focus	Number of studies[a]	Examples from studies
Improve students' academic skills	24	• Increase students' reading levels (Wallenstein, 2018). • Increase the number of students passing mathematics courses with a C or better (Rohanna, 2018).
Improve teacher practice	15	• Improve reading instruction (Taylor & Gordon, 2014). • Increase rigor of mathematics instruction (Park, 2019).
Improve students' social-emotional or behavioral skills	7	• Promote positive student behavior (Wilcox et al., 2017). • Develop student ownership and responsibility (Cannata et al., 2016).
Increase students' access to postsecondary opportunities	3	• Increase students' college attendance (Redding & Viano, 2018). • Foster students' postsecondary readiness (Wallenstein, 2018).
Improve parent and family engagement	2	• Promote family and community engagement (Wallenstein, 2018). • Promote transparent communication with parents of students who are failing courses (Wilcox et al., 2017).
Other	4	• Improve the professional development system (Edwards et al., 2015). • Improve the teacher feedback process (Hannan et al., 2015).

[a] Some studies were classified as having multiple foci for their CI initiative.

Who Were the Members of the Continuous Improvement Teams?

We also were interested in the roles of CI team members and if members represented different organizational levels (e.g., school vs. district level). Thirty-two of the 34 studies described the composition of CI teams (see Exhibit 12). CI teams most commonly included teachers and school-level administrators. Twenty-two studies described CI teams consisting of only school-level staff (e.g., teachers, guidance counselors, instructional coaches) and school-level administrators. Eleven studies described CI teams consisting of school-level staff, school-level administrators, and district personnel.

Exhibit 12. Teachers Were the Most Common Members of CI Teams in Studies

Role	Number of studies[a]
Teachers	25
School-level administrators	23
District personnel	13
Researchers	11
Instructional coaches or specialists	9
CI coaches or facilitators	7
Parents	5
University faculty	4
Guidance counselors	3
Other (classroom aide, school psychologist, student, support provider)	12

[a] Some studies were classified as having multiple roles represented on their CI teams.

CI Team Membership Selection Process

Engaging practitioners meaningfully in the CI work is a key component of the process. Eight studies (see Exhibit 13) detailed how all or some members of the CI team were selected to participate in the effort.

Exhibit 13. About One Fourth of Studies Described How All or Some Members of the CI Team Were Selected to Participate in the CI Effort

CI team selection process	Number of studies	Example from studies
Team members selected to be part of the team because of relevant experience or expertise	4	A science teacher was selected to be on the CI team based on experience with the Next Generation Science Standards (Sinapuelas et al., 2019).
Team members required to be part of the team	3	All reading teachers, all literacy coaches, and the assistant principals responsible for reading at each high school (Taylor & Gordon, 2014).
Team members volunteered to be part of the team	1	Four high school science teachers volunteered to participate in the collaborative teacher inquiry process (Ermeling, 2010).

Interventions, Number of Cycles, and Processes Used to Implement Continuous Improvement

What Types of Interventions Were Teams Trying to Implement?

The literature on continuous improvement (CI) in industry and health focuses on what might be called "process" improvements: identifying flaws in regular organizational routines that generate poor or varied performance (Langley et al., 2009). CI work is expected to begin by identifying a process or routine that is the root cause of the performance problem that will become the focus of improvement work. Once the root cause is uncovered, then attention turns to seeking a potential change idea to test.

Given its roots in process improvement, one might hypothesize that CI would be better suited to interventions or change ideas focused on concrete organizational routines—for example, the processes used to encourage and monitor students' completion of applications for college and financial aid—than to interventions focused on staff knowledge and skills—for example, teacher knowledge of typical student misconceptions about mathematical ideas.

Eighteen of the 34 studies described the intervention or interventions tested using CI methods. We classified the interventions based on a Change Idea Taxonomy developed by the Bill & Melinda Gates Foundation. The taxonomy includes five Level 1 categories: ambitious instruction, collaborative and knowledgeable teams, school systems and policies, supportive environment, and involved families and communities. Each Level 1 category contains more detailed Level 2 categories. The interventions described in these 18 studies fell in three of the five Level 1 categories—ambitious instruction, collaborative and knowledgeable teams, and supportive environment—and seven Level 2 categories (see Exhibit 14).

Eleven of the 18 studies described interventions in the ambitious instruction domain. Most of the remaining interventions focused on the supportive environment domain. It is hard to assess the extent to which the tested interventions focused on process improvements; however, one study described an intervention focused on process improvements within the college access and affordability domain. For example, the school team created and administered individualized college ready packets to increase students' awareness of their college eligibility (Aguilar et al., 2017).

Exhibit 14. About Half of Studies Described the Interventions Tested Using CI Methods

Level 1	Level 2	Number of studies[a]	Example from studies
Ambitious instruction	Instructional planning	6	Teachers designed a lesson that targeted a particular student misconception or deepened understanding of a particular mathematical idea and developed a hypothesis about anticipated student responses to the tasks provided in the lesson (Liang et al., 2019).
Ambitious instruction	Positive classroom environment	1	Teachers create a collaborative classroom culture in which all student voices are valued; for example, grouping students strategically to allow for equitable and accountable discourse (Park, 2019).
Ambitious instruction	Student-centered teaching	5	Mathematics teachers start class by implementing a 10-minute "opener" with student-to-student discourse related to rich tasks (Gallagher et al., 2019).
Collaborative and knowledgeable teams	Staff development and support	2	School teams created a teacher feedback process designed to support beginning teachers. Teams used improvement science methods to enhance, refine, and integrate the process into their existing systems (Hannan et al., 2015).
Supportive environment	College access and affordability	1	The school team created individualized "I Am Ready" packets for every senior who qualified to apply to California State University and University of California campuses to increase students' awareness of their college eligibility (Aguilar et al., 2017).
Supportive environment	Student agency, identity, and mindsets	3	Although each school implemented practices in somewhat distinct ways, four central practices were common across schools to foster student ownership and responsibility. These components were all taught during advisory periods. First, teachers taught students about growth mindsets. Second, teachers adopted practices that helped students monitor their grades. Third, students set goals of how they would reach academic targets. Finally, teachers worked with students to create problem-solving strategies for how they would reach these goals (Cannata et al., 2016).
Supportive environment	Supportive relationships and community	2	Personalization for academic and social-emotional learning is a systemic schoolwide and research-based approach in which administrators, guidance counselors, and teachers intentionally and deliberately attended to students' academic, social- emotional, and behavioral needs (Harrison et al., 2019; Tichnor-Wagner et al., 2017).

[a] Some studies discussed multiple interventions that fell into more than one Level 1 or Level 2 category.

How Many Cycles Did Continuous Improvement Teams Implement?

A key underlying premise of CI is that improvement requires an iterative inquiry process, in which ideas are perfected through repeated feedback. Thus, the number of cycles conducted might be viewed as one indicator of the quality with which CI has been implemented. Although inquiry cycles are at the "core" of CI, only 10 of the 34 studies reported on the number of cycles that CI teams implemented (see Exhibit 15).

Four of these 10 studies reported on the number of cycles implemented by a single team of practitioners. For example, Aggarwal and Lynn (2012) described a CI effort involving two cycles; the intervention was tested during one semester and then revised and tested again the following semester. The remaining six studies described inquiry cycles implemented by several teams. For example, Hannan et al. (2015) followed 10 schools for 3 years and reported that schools varied in the number of PDSA cycles implemented during that period; one school implemented 22 cycles, whereas another implemented eight.

Eight of the 10 studies took place over a 1-year period; one extended for 2 years (Liang et al., 2019) and one for 3 years (Wallenstein, 2018).

Exhibit 15. Ten Studies Discussed the Number of Cycles Implemented During the Study Period

Study	Study period	Number of teams	Number of cycles implemented (range)	Length of cycle
Aggarwal & Lynn (2012)	1 year	1	2	Semester long
Ell & Meissel (2011)	1 year	Multiple	2	Unknown
Ermeling (2010)	1 year	1	2	Unknown
Hannan et al. (2015)	1 year	Multiple	Varied across teams (ranged from 2 to 22)	Varied across teams
Liang et al. (2019)	2 years	1	2	Unknown
Park (2019)	1 year	Multiple	Varied across teams (range unknown)	Varied across teams
Rohanna (2018)	1 year	Multiple	Varied across teams (ranged from 9 to 22)	Varied across teams
Sinapuelas et al. (2019)	1 year	1	3	Unknown
Tichnor-Wagner et al. (2017)	1 year	Multiple	3	Unknown
Wallenstein (2018)	3 years	Multiple	Varied across teams (range unknown)	Varied across teams

What Processes Did Teams Use to Implement Continuous Improvement?

CI involves a complex set of routines. CI teams must learn how to identify a problem to focus on, determine potential root causes of the problem to test, select change ideas that might address the root causes, and engage in iterative cycles to test the change ideas. In most of the CI work that was the focus of the studies we reviewed, the CI teams were new to the CI process, and teams relied on a number of processes to build their knowledge and skill with respect to CI.

All 34 studies discussed the processes that teams used to implement CI—for example, the processes used to learn about inquiry cycles and put them into place. The most commonly discussed processes were the following:

- Collaborative learning opportunities (30 studies)
- Professional development or learning sessions (26 studies)
- Use of instruments (25 studies)
- Coaching (19 studies)
- Observation (9 studies)

Twenty-eight of the 34 studies described using more than one these processes, and these processes often overlapped with each other. For example, during a professional development session, practitioners were trained on how to use a measurement instrument.

Collaborative Learning Opportunities

Thirty studies described CI efforts in which team members engaged in collective learning through common planning time and/or by participating in a professional learning community or networked improvement community. These opportunities allowed teams to focus on an element of the inquiry process, such as examining student data after implementing an intervention. For example, Rohanna (2017) described school teams that participated in five group meetings over the course of one year to discuss improvement work.

Professional Development or Learning Sessions

Twenty-six studies indicated that CI efforts were supported by professional development or other learning sessions led by outside facilitators or coaches. Typically, the purpose of these sessions was to help teams learn about the elements of CI or improvement science; become trained on tools, such as rubrics or assessments related to CI; or learn about an evidence-based teaching strategy or intervention. For example, Gallagher and Cottingham (2019) described a series of workshops, facilitated by external experts, that were designed to help minimally experienced district leaders with implementing CI methods to lead and facilitate CI projects within their own district.

Use of Instruments

Twenty-five studies discussed how CI team members used specific instruments to conduct PDSA cycles or other parts of the CI process (e.g., root-cause analysis). Instruments that practitioners used to implement CI included rubrics, scoring guides, data protocols, or PDSA templates. These instruments were generally designed to help teams implement aspects of CI work with fidelity. For example, Taylor and Gordon (2014) described school teams that created standards-based assessments together to ensure common expectations of students, allow comparison of the results, and provide common data for teacher reflection during the analysis process.

Coaching

Nineteen studies discussed the use of coaching to implement CI efforts. We coded studies as including coaching components if internal or external professionals provided technical assistance, mentoring, and/or feedback to an individual practitioner or a whole team as part of the CI effort. In general, the goal of coaching was to increase team members' understanding of and ability to implement elements of the CI process. For example, Wallenstein (2018) discussed how all school teams were supported by an experienced educator who helped guide them through the CI process and then communicate learnings from their work to other teams. In other studies, coaches provided content expertise related to the intervention being tested.

Observation

Nine studies discussed the use of observation to support CI efforts. We coded studies as including an observation component if CI team members observed other practitioners as part of the CI process. For example, Ray (2014) discussed how a district developed an assessment tool to help district leaders determine how well schools were implementing CI efforts and identify best practices. The district team then conducted school site visits in which they observed classrooms and departments, conducted stakeholder focus groups, and reviewed relevant documentation. After the site visits were completed, the district team gave the school feedback to enhance their improvement efforts.

Outcomes of Continuous Improvement

What Types of Data Did the CI Team Members Collect?

Data collection and analysis are key elements of the continuous improvement (CI) process. CI teams are expected to collect data to help identify the root cause of the chosen problem, and also before, during, and after conducting inquiry cycles to determine if a change idea has improved the intended outcomes.

Twenty-five of the 34 studies explicitly described the types of data that the teams drew on in their CI work. Student academic data were the most common data source used by the CI teams, followed by practitioner observations and practitioner survey data (see Exhibit 16).

Exhibit 16. Student Academic Data Were the Most Common Source Collected by CI Teams

Data type	Number of studies that listed data type[a]	Example(s) of data types from studies
Student academics	17	Standardized test scores, classroom assessments, and attendance data
Practitioner observations	10	Peer observations using a rubric, videotaping classroom sessions, and reviewing afterward
Practitioner surveys	7	Instructional surveys and professional development evaluation surveys
Practitioner interviews	4	Teacher interviews and information discussions with instructors
Student interviews	4	Empathy interviews
Student surveys	4	Course evaluations
Student activities	3	Engagement data, attendance data, and the number of students who set up counseling appointments
Artifact analyses	3	Student coursework and teacher notes
Other	4	Districtwide climate surveys and student behavior data

[a] Some studies discussed multiple data types.

Did Student Outcomes Improve as a Result of the Continuous Improvement Effort?

The goal of most CI work in education is to improve student outcomes. Most of the studies we reviewed were focused on student learning as the primary outcome, and thus a key question is whether the studies provided evidence on if student outcomes improved over the course of the CI effort—and, if so, whether the improvement can be attributed to CI.

Of the 22 studies that described a CI initiative focused on improving students' academic outcomes, 11 provided evidence on at least one academic outcome measure, and all 11 indicated that the CI initiative led to improvement in some or all of the measured academic outcomes. Exhibit 17 lists the academic outcomes measured in each study, along with the measure used, the time period of measurement, if there was a comparison group, and whether reported improvements were statically significant. The 11 studies varied in their research designs: three used a causal-comparative design (e.g., designs with an explicit comparison group), three used a longitudinal design (following the same cohort of students across time to measure growth), and six used a consecutive cross-section design (comparing consecutive cohorts of students).

Of the three studies that used a causal-comparative design, two reported positive, statistically significant results (Arbour et al., 2016; Brown, 2013). For example, Brown (2013) examined students' reading assessment data from nine schools that fully, partially, or did not implement the eight steps of a CI model during the 2011–12 school year. Compared with the 2010–11 reading assessment scores, the author found that the test results showed "practical and significant differences in achievement scores among those schools that partially and fully implemented the model compared to those schools that did not implement [the model]" (Brown, 2013, p. iv).

The second study that used a comparison group (Arbour et al., 2016) focused on a group of preschool educators from 40 centers, who participated in professional development training on instructional strategies. The educators from about half of the centers participated in the training only, whereas the educators from the other half participated in a learning collaboration that taught them continuous quality improvement (CQI) skills in addition to the training. The authors compared the language scores of students taught by both groups, using propensity weights to adjust for preexisting differences in centers. The study found that "children in classrooms whose teachers received professional development in instructional strategies plus CQI methods had larger increases in language skills (vocabulary, effect size 0.31, $p < 0.05$) than children whose teachers received instructional strategies alone" (Arbour et al., 2016, p. 4).

The third study that used a comparison group (Gallagher et al., 2019) reported a positive difference on student achievement favoring the CI schools but did not report on the statistical significance of the difference. The study reported that, in a 2-year period, Grades 4–8 students' proficiency in mathematics grew in schools that launched Local Improvement Teams (LITs) compared with schools that did not launch LITs. To calculate schools' aggregate growth percentile, the authors developed a growth metric to estimate how much more or less students in a school with a LIT were learning relative to peers with similar achievement and demographic characteristics in schools without LITs. However, the authors did not report whether the change was statistically significant: "In [schools that launched Local Improvement Teams in 2017–18], even though a majority of students are not yet achieving proficiency in mathematics, students are exhibiting above-average growth relative to other schools serving similar students" (Gallagher et al., 2019, p. 10).

The other eight studies did not use comparison groups to assess whether CI led to an improvement in academic outcomes and instead examined pre-post change in outcomes using longitudinal or consecutive cross-section research designs. Four of these studies reported positive, significant results for all or some academic outcomes. For example, Howley-Rowe (2000a) reported data on the improvement in test scores across 4 years. The authors indicated that third-grade total basic skills scores increased in Bending Knee, the participating school, potentially caused by the improvement initiative (Quest Project):

> In terms of student achievement, improvement trends were apparent at Bending Knee between the years of 1997 and 2000. For instance, third grade total basic skills increased consistently between 1997 and 2000. Although some of the growth is statistically significant, only third grade math and total basic skills score increases seem to possess any practical significance as indicated by moderate effect sizes. Quest may have influenced the improvements in achievement at Bending Knee; however, such growth is relatively limited. Given that these findings are nonetheless promising, the school may see further growth as staff continue to implement Quest-related initiatives. (Howley-Rowe, 2000a, p. vi)

Of the remaining studies without comparison groups, four reported positive results in student achievement, but those results were either not statistically significant, or the authors did not report on the statistical significance of the results. For example, Arnold (2010) compared third-grade students' Missouri Assessment Program (MAP) test scores from 2008—prior to the CI initiative—to third-grade students' scores in 2009, after the CI initiative was implemented. The author found no statistically significant change in students' MAP test scores during the 1-year study period but did find achievement gains using Study Island, an online benchmarking

assessment tool that is an indicator of how students will perform on the MAP tests, as described in the following passage:

> In the fall, 31% of third grade students scored either Below Basic or Basic [on the Study Island benchmarks], while 69% scored either Proficient or Advanced. However, from the third benchmark assessment in the spring, only 14% of the students scored either Below Basic or Basic, while 86% of the students scored either Proficient or Advanced. This indicates that between the fall and the spring when the first and third benchmarks were conducted with the whole grade level, there was a 17% increase in mastery of the yearlong third grade objectives. (Arnold, 2010, p. 93)

In another study, Howley-Rowe (2000b) compared scores of fourth-grade students in 1997 to scores of fourth-grade students in 1998 and 1999 and found that "the percentage of fourth grade students performing on the Tennessee Writing Assessment at least at the competent level increased from 44% in 1998 to 63% in 1999," but the author did not discuss the statistical significance of this result.

Exhibit 17. Eleven Studies Discussed Improvement in Student Academic Outcomes

Study	Academic outcome	Measure(s)	Time period of measurement (measure)	Comparison group(s)	Study design	Results and statistical significance
Aggarwal & Lynn (2012)	Students' problem-solving skills	Course project grade	1 year	No	Pre-post, consecutive cross-section	+[a]
Arbour et al. (2016)	Preschool (4- and 5-year-olds) language scores	Woodcock-Munoz Language Survey	2 years	Yes	Causal-comparative with controls for premeasures	+ (mixed)[b]
Arnold (2010)	Grade 3 communication arts scores	Missouri Assessment Program (MAP), Study Island online benchmarking assessment	2 years (MAP); 1 year (Study Island)	No	Pre-post, consecutive cross-section and longitudinal	ns[c]
Brown (2013)	Grades 4, 8, and 10 reading/English language arts scores	Tennessee Comprehensive Assessment	1 year	Yes	Causal-comparative with controls for prior achievement	+ (mixed)[b]
Ell & Meissel (2011)	Grades 3–8 mathematics scores	Basic facts (math) assessments	<1 year	No	Pre-post, longitudinal	n/a[d]
Gallagher et al. (2019)	Grades 4–8 mathematics achievement and growth scores	Smarter Balanced Assessment Consortium (SBAC) test	2 years	Yes	Causal-comparative with controls for prior achievement	n/a[d]
Howley-Rowe (2000a)	Grades 3, 4, and 5 reading, mathematics, and language skills scores	Stanford Achievement Test	4 years (Grade 3 scores); 3 years (Grade 4 scores); 2 years (Grade 5 scores)	No	Pre-post, consecutive cross-section	+ (mixed)[b]

Study	Academic outcome	Measure(s)	Time period of measurement (measure)	Comparison group(s)	Study design	Results and statistical significance
Howley-Rowe (2000b)	Grade 4 writing scores	Tennessee Writing Assessment	3 years	No	Pre-post, consecutive cross section	n/a[d]
Ray (2014)	Grades 5 and 6 reading, mathematics, and science scores	Texas Assessment of Knowledge and Skills (TAKS); Brigance Comprehensive Inventory of Basic Skills Early Childhood Assessment; end-of-year benchmark scores	8 years (TAKS); 5 years (Brigance); 4 years (end-of-year benchmark scores)	No	Pre-post, consecutive cross section	+ (mixed)[b]
Taylor & Gordon (2014)	Grades 9–12 reading scores	Florida Comprehensive Assessment Test (FCAT)	2 years	No	Consecutive cross-section	n/a[d]
Walters (2012)	Grades 8–12 math achievement	TAKS	4 years	No	Pre-post, longitudinal	+ (mixed)[b]

[a] + indicates a positive and significant relationship for all outcomes. [b] + (mixed) indicates a positive and significant relationship for some outcomes. [c] ns: The study did not detect a significant relationship for any outcome. [d] n/a: The study did not discuss significance of any outcome.

Supportive Conditions and Barriers of Continuous Improvement

What Conditions Supported the Implementation of Continuous Improvement?

Developing and sustaining a culture of improvement within a school, district, or university involves multiple stakeholders. The work of continuous improvement (CI) can rarely be carried out in isolation, and team members require time to communicate with each other throughout the PDSA process. Thus, we examined the evidence the studies collected and analyzed about supports for collaboration, time to meet, and other factors supporting the implementation of CI.

Twenty-six of the 34 studies reported on at least some factors associated with implementation of CI. Most studies that focused on enabling conditions drew conclusions based on practitioner perceptions of factors that facilitated implementation. The authors generally did not define well-implemented CI or classify schools or other organizations as high or low implementers. Also, most studies did not explicitly define levels of support, although in some cases authors described leaders or others who were more or less supportive. Only one study directly categorized teams within high-, partial-, or low-implementing categories (Brown, 2013).

The conditions that emerged in the studies can be classified in two groups (see Exhibit 18): organizational leadership (supportive leadership, shared leadership, and organizational coherence) and opportunities for professional growth (quality professional development opportunities for collaboration and reflection, and respectful and trusting professional relationships).

Exhibit 18. Conditions That Supported Implementation of Continuous Improvement in Studies

Condition category	Conditions that supported implementation (number of studies[a])
Organizational leadership	• Supportive leadership (17) • Shared leadership (10) • Organizational coherence (10)
Opportunities for professional growth	• Quality professional development (16) • Opportunities for collaboration and reflection (11) • Respectful and trusting professional relationships (11)

[a] Some studies discussed multiple conditions that supported implementation.

Supportive Leadership

Seventeen of the 34 studies described supportive school and/or district leadership as a factor facilitating implementation. Support from administrators could include establishing internal structures to support the CI effort and/or actively engaging in the CI process. For example, Gallagher et al. (2019) discussed how districts provided local improvement teams with additional resources to collect and analyze data required for the CI work. Ermeling (2010) described how the assistant principal freed up time for the CI team to meet by reducing their responsibilities during a standardized testing day. Finally, Rohanna (2018) noted the importance of supportive leadership in the CI process as follows:

> When establishing the school work structures, administrators play a significant role. In the schools w[h]ere administrators were more involved and supportive of the network, it was easier for the hub to establish consistent meetings with teachers. In the cases where administrators were either not fully involved, or did not value the process, it was more difficult to schedule regular meetings. Supportive administrators are needed to carve out the time and space for their teachers to meet. Without their cooperation and assistance, it will be difficult to [sic] for their teachers to implement improvement science activities. (p. 140)

Shared Leadership

Ten of the 34 studies discussed how shared leadership between practitioners supported the success of a CI effort. Sharing leadership across CI team members and/or levels of a school or district increased practitioner buy-in to the CI process. For example, Kennedy and Gallagher (2019) discussed how teachers and school leaders worked together on the CI effort and, during interviews with the teacher members, "teachers described working harder than ever through the improvement science process, but also feeling joy and fulfillment, which stands in stark contrast to many initiatives [that] administrators ask teachers to undertake" (p. 10). Redding and Viano (2018) described an innovation design effort that involved both district and school personnel; they concluded that including teachers in the design process helped create an innovation that fit into teachers' instructional context, thus generating teacher buy-in for the CI effort:

> Allowing for teacher self-determination in the innovation design and implementation helped to garner a high level of teacher buy-in to the innovation. Compared with externally developed reforms, the innovation was less challenging to teacher autonomy and was customized to fit the needs of their students. These conditions led to high levels of teacher ownership over the innovation. Yet, in the process, teacher leaders grounded the innovation in preexisting and easy-to-implement practices that did not require significant investment from teachers to adopt. (p. 1)

Organizational Coherence

Ten of the 34 studies described organizational coherence as a supportive condition for CI. We defined organizational coherence as the alignment of CI work with the school, district, or university's vision, goals, and/or systems. Supportive leaders helped create this alignment, which, in turn, generated buy-in among practitioners. For example, Wallenstein (2018) examined the New York City Department of Education's effort to build schools' capacity for improvement through the Learning Partners Program (LPP), which required school teams to conduct inquiry cycles known as cycles of learning. The author found that "schools that found alignment between their overall school improvement goals and the work of the program were more likely to sustain focus on LPP" (Wallenstein, 2018, p. 203). Wilcox et al. (2017) also described the importance of organizational coherence in their discussion of the Department of Defense Education Activity's (DoDEA) effort to become a CI organization:

> While the concept of continuous improvement is simple, the implementation is quite complex. Even small-scale changes to a system have ripple effects and unforeseen challenges, so change to a system as large as DoDEA is certainly a complicated venture. Although implementation of continuous improvement is still in early stages, many lessons have been learned. Having a common goal (i.e., purpose) to rally around is essential. It acts as a motivator and director for all change and assists with clear communication and consistency across regions, districts, and schools. The strategic plans also help to communicate that direction and purpose by providing structure, while retaining enough flexibility to allow for creation and negotiation of details and plans for achieving the years' goals. (p. 15)

Quality Professional Development

Sixteen of the 34 studies mentioned professional development as a key support. For example, Gallagher and Cottingham (2019) discussed the importance of providing scaffolded professional development sessions to train practitioners on how to use CI-specific tools as they engaged in each step of the process: "This [scaffolded] approach often focuses on teaching educators specific tools (e.g., fishbone diagram, Plan Do Study Act form) designed to help novices successfully execute small steps that cumulatively support the execution of an important component of continuous improvement" (p. 8). Hardy (2016) also found that CI team members valued having an instructional expert provide professional development to them throughout the year. The author concluded that these sessions helped improve the teachers' practice and supported their CI work:

> Teachers were stimulated by exposure to academic literature (the Timperley article), engaging with one another around their teaching practice, and active dialogue development on student assessment data (and analyses of data in general), and these all helped to give a sense of confidence to these teachers about their practice. (Hardy, 2016, p. 15)

Opportunities for Collaboration and Reflection

Eleven of the 34 studies discussed providing opportunities for collaboration and reflection as a supportive condition for CI. For example, Ray (2014) described how principals from five schools met regularly to conduct strategic planning, which included examining the implementation of strategic objectives, reviewing documentation, identifying areas of improvement, and showcasing best practices. The author found that this review process was "an impetus for continuous improvement as evidenced through positive student performance over time" (Ray, 2014, p. 5). Several studies also noted how providing time for teachers to reflect and discuss their practice helped them develop an inquiry-based CI effort instead of a compliance-based one. Ell and Meissel (2011) examined how collaborative meetings in which teachers discussed student data and their own instructional practice promoted collective responsibility:

> The schools engaged in data collection, received feedback in the form of analysed cluster data and made action plans (including planned professional development in three schools) based on the results. For the teachers that attended cluster meetings[,] this process opened up discussion, reduced blame, aided the emergence of a shared language and understanding about the mathematics underachievement problem, encouraged collective responsibility for change and began to raise questions about teacher effects and the relationship between teaching and learning. (p. 182)

Respectful and Trusting Professional Relationships

Eleven of the 34 studies described respectful and trusting professional relationships as key to the successful implementation of CI. Several studies discussed how collaborative learning opportunities fostered trust among practitioners. Wallenstein (2018) illustrated how teams visiting each other's schools and school team meetings helped practitioners form trusting relationships, which, in turn, led to increased innovation:

> Observations of interschool visits and school team meetings demonstrated that schools and networks with trusting relationships and normative dispositions engaged in open discussions of challenges and provided critical feedback, which led to concrete ideas for changes to practice. Developing these conditions was also crucial for overcoming issues related to discomfort with the host and partner role designations. (p. 197)

Other study findings exemplify how the preexistence of trusting professional relationships support a CI effort. For example, Park (2019) discussed how the Long Beach Unified School District built and sustained a culture of continuous learning, noting how its staff "highlight trusting relationships fostered by decades of consistent leadership and staff as a key feature that signals peoples' investment in the organization and in one another" (p. 2).

What Were the Barriers to Implementing Continuous Improvement?

Twenty-one of the 34 studies discussed one or more barriers to the implementation of CI work. The most common barriers to improvement were related to lack of staff capacity or expertise and time constraints. Other barriers included staffing issues, lack of alignment between school or district initiatives, lack of consistency in implementation, and lack of communication or trust among members of the CI team.

We defined the lack of capacity or expertise as staff having limited knowledge, understanding, buy-in, or experience implementing CI. For example, in Cannata et al. (2016), CI teams initially resisted using data to measure the success of an intervention, which limited their ability to successfully implement an intervention designed to improve student outcomes. Similarly, Tichnor-Wagner et al. (2017) discussed how practitioners implementing PDSA cycles in their districts had a mixed response to the CI initiative because they felt it was disconnected from their daily work; the authors noted that this lack of buy-in could have been caused by a lack of fiscal resources for creating a proper infrastructure to support the initiative.

The barrier of time was most commonly portrayed as staff not having enough time to implement CI processes. For example, Shearer (2018) found district teams, composed of members involved in an improvement initiative that occurred in two academic years, said that districts need more than a few years to determine whether the approach benefits students. Similarly, Cannata et al. (2016) noted that there are contractual limits on teachers' time, which ultimately place time restrictions on how often staff could meet to discuss the progress of their improvement work.

Diffusion of Continuous Improvement Practices

Did the Lessons Learned From Implementing an Intervention Using Continuous Improvement Spread Beyond the Team?

One goal of continuous improvement (CI) is to share what is learned beyond the CI team, including knowledge about interventions that were found to be effective as well as knowledge about change ideas that seem promising but did not succeed. Some conceptualizations of CI explicitly involve linking CI teams in social networks to foster sharing, with regular meetings and routines to support sharing (e.g., Bryk et al., 2015). We did not restrict our review of CI to initiatives that were explicitly situated in networks, although many were. Even among teams not explicitly joined in networks, team members may share what they learn through other networks of which they are a part of (e.g., grade-level or subject-matter teams).

Fourteen of the 34 studies discussed diffusion, the spread of ideas generated during the CI work, beyond the initial CI team. We classified these studies in two ways. First, we noted whether the diffusion occurred within the schools, districts, or universities involved in the CI process or extended beyond these organizational units (see Exhibit 19). Eleven studies described how practitioners shared ideas within the organizations involved in the change effort; three studies discussed how practitioners shared ideas with other organizations involved in the CI effort (e.g., organizations linked in a network); and three studies discussed how practitioners shared ideas with organizations not involved in the CI effort.

In addition, we were interested in understanding the intervention or process diffused in each study. Some studies examined the spread of the intervention that was tested through CI, some examined the spread of the methods used to carry out the inquiry process, and some examined the spread of lessons learned during the inquiry process (e.g., modifications to the intervention).

Exhibit 19. Fourteen Studies Described Diffusion Efforts That Took Place During and After the CI Effort

Setting	Number of studies[a]	Example of what was diffused
Within the school/district/ university engaged in the CI effort	11	• **Intervention**. Teachers piloted an intervention in one Grade 3, Grade 4, and Grade 5 classroom as part of the CI initiative. During the following school year, the intervention was extended to all kindergarten through Grade 5 classrooms (Howley-Rowe, 2000b). • **Intervention**. Three high schools implemented an intervention that was eventually scaled up to other high schools across the district (Tichnor-Wagner et al., 2017). • **Method**. After one CI effort at a school concluded in the 2003–04 school year, the entire faculty was organized into working groups the next fall (2005) to implement the same inquiry process (Ermeling, 2010). • **Lessons learned**. Five school teams from the same district shared results of inquiry cycles (Rohanna, 2018). • **Lessons learned**. School teams in the same district engaged in monthly interschool site visits in which teams exchanged ideas and engaged in learning activities (Wallenstein, 2018).
With other schools/districts/ universities engaged in the CI effort	3	• **Intervention**. Schools adopted interventions from other schools in the Quest Network, a network consisting of 17 schools across Kentucky, Tennessee, Virginia, and West Virginia (Howley-Rowe, 2000a). • **Lessons learned**. Teachers from 40 separate preschools in three municipalities in Chile presented their tests of change to one another and shared learning across schools (Arbour et al., 2016).
With other schools/districts/ universities not involved in the CI effort	3	• **Intervention**. The intervention tested during the CI initiative at one university was adopted by six universities not part of the original CI effort (Sinapuelas et al., 2019). • **Intervention**. A version of the teacher inquiry intervention tested during the CI effort was implemented at more than 100 secondary schools in California and other parts of the United States (Ermeling, 2010).

[a] Some studies discussed diffusion efforts that fell into multiple settings.

Summary and Conclusions

We located 34 empirical studies that focused on the implementation of continuous improvement (CI) in education settings. The studies vary considerably in depth, types of data collected, analyses reported, and conclusions. Taken together, one thing they suggest is that although there is a good deal of consensus on the conceptualization of CI and a vast literature on the rationale and role of CI, there are relatively few empirical studies of the implementation of CI and little consensus on how it should be studied. Empirical research on the implementation of CI is clearly in its infancy; thus, the conclusions that can be drawn about the typical level of implementation of CI, the outcomes of CI, and factors that support or limit implementation must be viewed as tentative.

The following main themes emerged from our analysis pertaining to each topic we examined in the review.

- **The ways the studies defined CI**. The boundary between studies of CI and other approaches to reform (e.g., professional learning communities, data use, research-practice partnerships, and implementation science) are not clear-cut. A review of the literature including studies in these related areas would be vast; thus, we focused on studies that explicitly have roots in the Model of Improvement or similar conceptualizations. Identifying studies that "squarely" fall within this domain is not straightforward. The phrase "continuous improvement" often is used in a generic sense, to mean "efforts to improve," not the application of inquiry cycles.
- **The research designs employed in the studies**. Almost all of the reviewed studies are cases, and many focused on relatively few teams. Fourteen studies focused on just one school or other entity implementing CI, and another 12 focused on fewer than five entities. Studies comparing schools, districts, or universities that have engaged in CI with those that have not are quite rare. Only three of the reviewed studies involved the explicit comparison of teams that adopted CI with those that had not adopted CI. Furthermore, because CI often involves a partnership between practitioners and researchers, many CI studies are authored by participants. In more than half of the studies we reviewed, the authors were participants in the CI process. This gave the authors close access to data, but, at the same time, it might be useful to encourage more "third-party" studies.
- **The duration, setting, and focus of CI initiatives**. Most studies examined CI efforts that were relatively short; over half of the studies reported on a CI effort that was in its first or second year. One implication is that the practitioners and other participants in the CI process were learning to implement CI at the same time they were trying to improve the chosen intervention. Research on what might be called the "mature" use of CI is very

limited. In addition, because most studies focused on the first few years of CI, it is not yet possible to draw conclusions on whether the level of implementation typically improves with experience or how long it takes to learn to implement CI well.

- **The number of cycles used to implement CI**. Although iterative inquiry cycles are a fundamental component of CI, relatively few studies provided detailed data on the implementation of cycles. Of the 34 studies, only 10 studies reported on the number of cycles implemented, and those that did generally indicated that the number was 2 or 3. Furthermore, relatively few studies gave examples of cycles; thus, it is difficult to know how well they were implemented. Clearly, more attention should be given to research methods that can elicit valid information on the nature and implementation of inquiry cycles.
- **The outcomes of CI**. Only 11 studies measured student achievement outcomes. All 11 studies reported at least some positive results, but only three studies reported on explicit comparisons of achievement in schools that implemented CI with those that did not. Of these three studies, two reported statistically significant differences favoring the CI group, and the third indicated positive results without reporting statistical significance. The remaining eight studies examined pre-post improvement in schools that implemented CI. Four of the eight studies reported positive, significant results for all or some academic outcomes. Taken together, the studies appear to provide some support for the hypothesis that implementing CI can lead to improved outcomes for students, but the research designs do not permit strong causal conclusions.
- **Supportive conditions and barriers of CI**. Most of the studies we reviewed drew conclusions about factors facilitating and constraining the implementation of CI, basing these conclusions on the perceptions of participants in the CI process. Only one study explicitly compared conditions associated with high- and low-implementing CI teams. We classified the authors' conclusions into six conditions that supported CI implementation: supportive leadership, shared leadership, organizational coherence, quality professional development, the opportunity for collaboration and reflection, and respectful and trusting professional relationships. Participants reported that lack of expertise and insufficient time were barriers to implementation.
- **The diffusion of CI**. One goal of implementing CI is to develop and improve change ideas that can be spread to other teachers, schools, districts, and institutions. Less than half of the 34 studies discussed diffusion. Of these, most discussed diffusion that occurred within the organization that engaged in CI (e.g., school). We found that studies commonly discussed the spread of interventions used and lessons learned about the CI process.

Overall, our results suggest that much remains to be learned about CI in practice. The available empirical studies provide what might be called an "existence proof," indicating that CI teams can implement inquiry cycles in the ways imagined in the conceptual literature, especially under the right conditions—for example, with supportive leadership, professional development, and opportunities for collaboration. And the studies provided suggestive evidence that implementation of CI might be associated with improvement in outcomes.

But there is much we do not know. Few studies provide examples of CI work extending for more than a few cycles or years; thus, we know very little about whether it is possible to sustain CI work over time or whether extended CI would lead to more substantial improvements in outcomes.

We also have very little information on whether CI is easier to implement or more effective for some types of interventions or change ideas than others. Although the literature provides evidence that inquiry cycles can be carried out on a variety of topics (e.g., instruction, social-emotional learning, and college access), there are too few studies on each to draw conclusions about the relative efficacy of CI for different topics. We also do not have strong evidence on the quality of CI implementation. Although we know that CI teams were able to put inquiry cycles in place, we have much less knowledge about the quality of the work, especially the nature of the evidence that teams drew on in the "study" phase of the PDSA cycle and the ways in which teams drew conclusions in the "act" phase. Finally, although there is some evidence that engaging in CI may be associated with an improvement in outcomes, the available evidence is based on very weak designs—with only three studies based on an explicit "no CI" comparison group and no studies based on random assignment.

Taken together, the 34 studies support the premise that CI is a promising avenue for reform, but more work is necessary to assemble evidence on the use of iterative inquiry cycles to improve student outcomes. One challenge in doing this is that CI is by its nature a local effort embedded in practice, not an approach that involves researchers testing the efficacy of a standardized intervention. To build an evidence base for this improvement approach, at least some CI efforts need to be undertaken within a systematic research framework so that the focus is not only on local learning by the CI team but also on more general learning about CI by researchers, practitioners, and policymakers.

References

*Adams,[5] D. (2013). *Creating a new learning community: A case study of school leaders' perceptions of the Baldrige Criteria for Performance Excellence and their implementation in a suburban school district* [Unpublished doctoral dissertation]. Northern Illinois University.

*Aggarwal, A. K., & Lynn, S. A. (2012). Using continuous improvement to enhance an online course. *Decision Sciences Journal of Innovative Education*, *10*(1), 25–48. https://doi.org/10.1111/j.1540-4609.2011.00331.x

*Aguilar, J., Nayfack, M., & Bush-Mecenas, S. (2017). *Exploring improvement science in education: Promoting college access in Fresno Unified School District*. Stanford Graduate School of Education, Policy Analysis for California Education. https://edpolicyinca.org/sites/default/files/FUSD-continuous-improvement.pdf

*Arbour, M., Yoshikawa, H., Atwood, S., Duran Mellado, F. R., Godoy Ossa, F., Trevino Villareal, E., & Snow, C. E. (2016). *Improving quality and child outcomes in early childhood education by redefining the role afforded to teachers in professional development: A continuous quality improvement learning collaborative among public preschools in Chile*. Society for Research on Educational Effectiveness. https://eric.ed.gov/?id=ED566990

*Arnold, A. J. (2010). *Exploring the use of the eight-step process in the area of third grade communication arts* [Unpublished doctoral dissertation]. Lindenwood University.

*Aumiller, B. E. (2008). *Implementation of the Baldrige education framework for performance excellence: A case study of one Midwestern school district leadership team's continuous improvement efforts* (Publication No. 3314727) [Doctoral dissertation, University of Illinois at Urbana-Champaign]. ProQuest.

Barletta, B., Comes, D., Perkal, J., Schumaker, R., Wallenstein, J., & Yang, B. (2018). *Networks for school improvement: A review of the literature*. Columbia University Center for Public Research and Leadership. https://cprl.law.columbia.edu/sites/default/files/content/CPRL-Gates%20Report-082318-FINAL.pdf

[5] All references preceded by an asterisk indicate that the study was included as part of this literature review.

Berwick, D. M. (1989). Continuous improvement as an ideal in health care. *New England Journal of Medicine*, *320*(1), 53–56. https://doi.org/10.1056/NEJM198901053200110

Best, J., & Dunlap, A. (2014). *Continuous improvement in schools and districts: Policy considerations*. McREL International. https://files.eric.ed.gov/fulltext/ED557599.pdf

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.

*Brown, J. V. (2013). *The continuous improvement model: A K–12 literacy focus* [Unpublished doctoral dissertation]. Lipscomb University.

Bryk, A. S. (2018, April 3–5). *Advancing quality in continuous improvement* [Keynote lecture]. Carnegie Foundation 2018 Summit on Improvement in Education, San Francisco, CA, United States. https://www.carnegiefoundation.org/resources/publications/advancing-quality-in-continuous-improvement/

Bryk, A. S., Gomez, L., & Grunow, A. (2011). Getting ideas into action: Building networked improvement communities in education. In M. Hallinan (Ed.), *Frontiers in Sociology of Education* (pp. 127–162). Springer.

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.

Bush-Mecenas, S., Christianson, K., Prado Tuma, A., Gahlon, G., Rico, M., & Ivy, T. (2020). *Interorganizational network effectiveness and network characteristics: A review of a decade of research literature* (Report No. PR-A242-1). RAND Corporation: Education and Labor.

*Cannata, M., Redding, C., & Rubin, M. (2016). *Continuous improvement in action: Educators' evidence use for school improvement*. Society for Research on Educational Effectiveness. https://eric.ed.gov/?id=ED567214

Coburn, C., & Russell, J. (2008). Getting the most out of professional learning communities and coaching: Promoting interactions that support instructional improvement. *Learning Policy Brief*, *1*(3), 1–5. https://www.sesp.northwestern.edu/docs/publications/7373301757c9ac9dd89f0.pdf

Coburn, C. E., Penuel, W. R., & Geil, K. E. (2013). *Practice partnerships: A strategy for leveraging research for educational improvement in school districts*. William T. Grant Foundation. https://eric.ed.gov/?id=ED568396

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

Chalmers, I., & Altman, D. G. (Eds.). (1995). *Systematic reviews*. BMJ Publishing.

Deming, W. E. (1986). *Out of the crisis*. Massachusetts Institute of Technology.

Eccles, M. P., & Mittman, B. S. (2006). Welcome to implementation science. *Implementation Science*, *1*, Article 1.

*Edwards, A. R., Sandoval, C., & McNamara, H. (2015). Designing for improvement in professional development for community college developmental mathematics faculty. *Journal of Teacher Education*, *66*(5), 466–481. https://doi.org/10.1177/0022487115602313

*Ell, F., & Meissel, K. (2011). Working collaboratively to improve the learning and teaching of mathematics in a rural New Zealand community. *Mathematics Education Research Journal*, 23(2), 169–187. https://doi.org/10.1007/s13394-011-0010-7

*Ermeling, B. A. (2010). Tracing the effects of teacher inquiry on classroom practice. *Teaching and Teacher Education*, *26*(3), 377–388. https://doi.org/10.1016/j.tate.2009.02.019

Fogarty International Center. (2021). *Implementation science news, resources and funding for global health researchers*. National Institutes of Health. https://www.fic.nih.gov/ResearchTopics/Pages/ImplementationScience.aspx

*Gallagher, H. A., & Cottingham, B. W. (2019). *Learning and practicing continuous improvement: Lessons from the CORE districts*. Stanford Graduate School of Education, Policy Analysis for California Education. https://edpolicyinca.org/sites/default/files/R_Gallagher_Oct19_0.pdf

*Gallagher, H. A., Gong, A., Hough, H. J., Kennedy, K., Allbright, T., & Daramola, E. J. (2019). *Engaging district and school leaders in continuous improvement: Lessons from the 2nd year of implementing the CORE Improvement Community*. Stanford Graduate School of Education, Policy Analysis for California Education. https://edpolicyinca.org/sites/default/files/CORE_Gallagher_Jan-2019.pdf

*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

*Hardy, I. (2016). In support of teachers' learning: Specifying and contextualising teacher inquiry as professional practice. *Asia-Pacific Journal of Teacher Education*, *44*(1), 4–19. https://eric.ed.gov/?id=EJ1086407

*Harrison, C., Wachen, J., Brown, S., & Cohen-Vogel, L. (2019). A view from within: Lessons learned from partnering for continuous improvement. *Teachers College Record*, *121*(9), 1–38. https://eric.ed.gov/?id=EJ1225420

*Howley-Rowe, C. (2000a). *Bending Knee Elementary: A case study of the Quest Network*. AEL. https://eric.ed.gov/?id=ED456507

*Howley-Rowe, C. (2000b). *Bowman Elementary: A case study of the Quest Network*. AEL. https://eric.ed.gov/?id=ED456506

*Huffman, D., & Kalnin, J. (2003). Collaborative inquiry to make data-based decisions in schools. *Teaching and Teacher Education*, *19*(6), 569–580. https://eric.ed.gov/?id=EJ680014

*Kaufman, E. K., Cash, C., Coartney, J., Ripley, D., Guy, T., Glenn, W., Mitra, S., & Anderson, J. C., II. (2019). Planning to create a culture of continuous improvement with the Department of Defense Education Activity. *Educational Planning*, *26*(4), 5–19. https://eric.ed.gov/?id=EJ1234914

*Kennedy, K., & Gallagher, H. A. (2019). *Leadership that supports continuous improvement: The case of Ayer Elementary*. Stanford Graduate School of Education, Policy Analysis for California Education. https://edpolicyinca.org/sites/default/files/R_Kennedy_Oct19_0.pdf

Kimmelman, P. L., & Kroeze, D. J. (2002). *Achieving world-class schools: Mastering school improvement using a genetic model*. Christopher-Gordon. https://eric.ed.gov/?id=ED478528

Langley, G. J., Moen, R. D., Nolan, K. M., Nolan, T. W., Norman, C. L., & Provost, L. P. (2009). *The improvement guide: A practical approach to enhancing organizational performance*. John Wiley & Sons.

Lewis, S. (2015). Qualitative inquiry and research design: Choosing among five approaches. *Health Promotion Practice, 16*(4), 473–475. https://doi.org/10.1177/1524839915580941

*Liang, S., Mira, R. V., Prasad, P. V., & Patterson, C. L. (2019). Improving our practice as mathematics teacher educators through teaching research. *International Journal for the Scholarship of Teaching and Learning*, 13(2), Article 12. https://eric.ed.gov/?id=EJ1218282

Miller, P. G., Curtis, A., Sønderlund, A., Day, A., & Droste, N. (2015). Effectiveness of interventions for convicted DUI offenders in reducing recidivism: A systematic review of the peer-reviewed scientific literature. *The American Journal of Drug and Alcohol Abuse*, 41(1), 16–29. https://doi.org/10.3109/00952990.2014.966199

Networks for School Improvement Initiative. (2019). *Continuous improvement core parameters*. Catalyst Ed. https://nsiexchange.org/wp-content/uploads/2019/04/CoP-Core-Parameters-Activity-Notes-Final-Version.pdf

Page, J., & Eadie, P. (2019). Coaching for continuous improvement in collaborative, interdisciplinary early childhood teams. *Australasian Journal of Early Childhood*, *44*(3), 270–284. https://doi.org/10.1177/1836939119855542

*Park, S., Hironaka, S., Carver, P., & Nordstrum, L. (2013). *Continuous improvement in education*. Carnegie Foundation for the Advancement of Teaching. https://www.carnegiefoundation.org/wp-content/uploads/2014/09/carnegie-foundation_continuous-improvement_2013.05.pdf

*Park, V. (2019). *Bridging the knowing-doing gap for continuous improvement: The case of Long Beach Unified School District*. Stanford Graduate School of Education, Policy Analysis for California Education. https://edpolicyinca.org/sites/default/files/R_Park_Oct19_1.pdf

Pate, A., Lerner, J. B., & Browning, A. (2012). *Beyond the numbers: Data use for continuous improvement of programs serving disconnected youth*. American Youth Policy Forum. https://eric.ed.gov/?id=ED538971

*Ray, J. L. (2014). *A multiple case study of a Baldrige-based peer review process in an urban school district* (Publication No. 3536718) [Doctoral dissertation, Sam Houston State University]. ProQuest.

*Redding, C., & Viano, S. (2018). Co-creating school innovations: Should self-determination be a component of school improvement? *Teachers College Record*, *120*(11), 1–32. https://eric.ed.gov/?id=EJ1187509

Regional Educational Laboratory Program West. (2017). *Introduction to improvement science*. U.S. Department of Education, Institute of Education Sciences. https://ies.ed.gov/ncee/edlabs/regions/west/Blogs/Details/2#improvement_science

*Rohanna, K. (2017). Breaking the "adopt, attack, abandon" cycle: A case for improvement science in K–12 education. *New Directions for Evaluation*, *153*, 65–77.

*Rohanna, K. (2018). *Preparing schools to successfully participate in networked improvement communities: A case study of year 1 of a math instructional network*. University of California–Los Angeles.

Sabelli, N. H., & Harris, C. J. (2015). The role of innovation in scaling up educational innovations. In C.-K. Looi & L. W. Teh (Eds.), *Scaling educational innovations* (pp. 13–30). Springer.

*Shearer, K. (2018). *Creating systems of continuous improvement: A phenomenological study of California districts engaged in transforming into learning organizations* [Unpublished doctoral dissertation]. Brandman University.

Shewhart, W. A. (1991). *Economic control of quality of manufactured product*. Van Nostrand Reinhold.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4–14. https://doi.org/10.3102/0013189X015002004

*Sinapuelas, M. L., Lardy, C., Korb, M. A., Bae, C. L., & DiStefano, R. (2019). Developing a three-dimensional view of science teaching: A tool to support preservice teacher discourse. *Journal of Science Teacher Education*, *30*(2), 101–121. https://doi.org/10.1080/1046560X.2018.1537059

Taylor, M. J., McNicholas, C., Nicolay, C., Darzi, A., Bell, D., & Reed, J. E. (2014). Systematic review of the application of the plan–do–study–act method to improve quality in healthcare. *BMJ Quality & Safety*, *23*(4), 290–298. https://doi.org/10.1136/bmjqs-2013-001862

*Taylor, R. T., & Gordon, W. R., II. (2014). 4 schools, 1 goal: University-district partnership nets results for struggling readers. *The Learning Professional*, *35*(3), 16–20. https://learningforward.org/wp-content/uploads/2014/06/4-schools-1-goals.pdf

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://eric.ed.gov/?id=EJ1157388

*Wallenstein, J. (2018). *Investing in school learning: The New York City Department of Education's Learning Partners Program*. Columbia University. https://academiccommons.columbia.edu/doi/10.7916/D8PP0HMD

*Walters, M. B. (2012). *The impact of the data teams process on student mathematics achievement* (Publication No. 3575036) [Doctoral dissertation, Texas Southern University]. ProQuest.

*Wilcox, K. C., Lawson, H. A., & Angelis, J. I. (2017). COMPASS-AIM: A university/P–12 partnership innovation for continuous improvement. *Peabody Journal of Education*, *92*(5), 649–674. https://doi.org/10.1080/0161956X.2017.1368654

Youn, S. J., Valentine, S. E., Patrick, K. A., Baldwin, M., Chablani-Medley, A., Aguilar Silvan, Y., Shtasel, D. L., & Marques, L. (2019). Practical solutions for sustaining long-term academic-community partnerships. *Psychotherapy*, *56*(1), 115–125. https://doi.org/10.1037/pst0000188

Appendix A. Literature Review Purpose and Methods

We designed the literature review to identify and code empirical studies that focused on the implementation of continuous improvement (CI) in education and health. To structure our review, we developed a set of four guiding questions (Qs).

Q1: What is known about CI?

- What are the "core features" of CI?
- What features are hypothesized to be associated with effectiveness?

Q2: What processes are used in the implementation of CI models, and how do those processes vary?

- What types of problems and interventions have entities using CI focused on in their work? In particular, to what extent have schools, districts, or other entities focused on instruction versus student behavior, such as attendance, college application, and so on?
- How has the implementation of CI been measured?
- Have studies used artifacts to measure the implementation of CI? If so, what types of artifacts have been used, and how have they been coded?

Q3: Who is the CI "team"? What roles do they play?

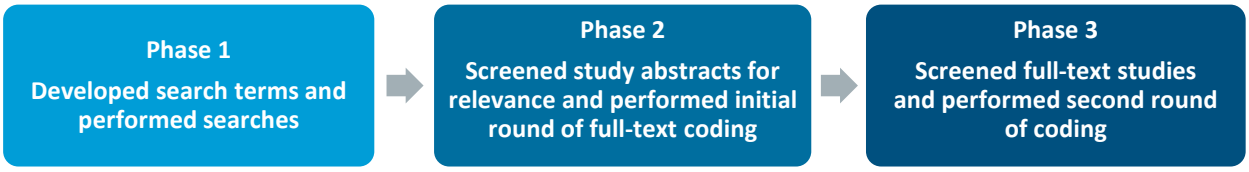
- Who has participated in CI teams (e.g., teachers; principals; other staff, such as counselors, students, and parents)?
- What supports have helped foster the implementation of CI (e.g., coaching, a supportive principal, planning time)?

Q4: How does CI diffuse?

- How have the practices developed and tested through CI diffused outside the CI team?

We based the search and screening process on the systematic literature process described by Chalmers and Altman (1995). The search and screening process involved three phases (see Exhibit A1).

Exhibit A1. Search, Screen, and Coding Phases



Phase 1. Developed Search Terms and Performed Searches

We began the review by developing search terms to drive the literature review. We established these search terms based on the guiding questions and then searched major academic databases for studies related to the terms. First, we created general search terms to capture the review's overarching topics of healthcare, education, and CI. We then developed more specific search terms based on the four questions (see Exhibit A2).

Exhibit A2. Search Terms by Guiding Question

Guiding question(s)	Search terms	
All Qs: healthcare and education	• Education • Health • Healthcare • Medicine	
All Qs: continuous improvement	• CI • Continuous improvement • Continuous quality improvement • Implementation science	• Improvement science • Networked improvement communities • NIC • Theory of practice improvement
Q1: What is known about CI?	• Concept • Case studies • Case study • Definition	• Description • Term • Terminology • Theory
Q2: What processes are used in the implementation of CI models, and how do those processes vary?	• Condition • Data driven • Disciplined inquiry • Driver diagram • Driver diagrams • Evaluation • Evidence-based practice • Evidence-based practices • Fishbone diagram • Intervention	• Inquiry circles • Inquiry cycles • Logic model • Logic models • Research-based practice • Research-based practices • Root causes • PDSA • Plan-Do-Study-Act

Guiding question(s)	Search terms	
Q3: Who is the CI "team"? What roles do they play?	• Implementation team • Network • Team member • Team members	
Q4: How does CI diffuse?	• Collaboration • Collective efficacy • Diffusion • Network	• Research-practice partnership • Stakeholder • Shared decision making

We performed the searches using three databases: ERIC, SocINDEX, and PsycINFO. All the databases were searched via the EBSCO interface. We limited our searches to sources published in English within the last 25 years (1995–2020). We performed four separate searches to find sources related to each question. During each search, we included guiding question–specific terms as well as general terms related to healthcare, education, and CI (as listed in Exhibit A2). For each search, the topics for the first two rows in Exhibit A2 and the row corresponding to one of the guiding questions were connected via a Boolean string AND the search terms within each topic were connected by a Boolean string OR used quotation marks to search for terms as necessary. For example, the Q1 search was constructed as follows:

> (CI OR "continuous improvement" OR "continuous quality improvement" OR "implementation science" OR "improvement science" . . .)
>
> AND
>
> (education OR health OR healthcare OR medicine)
>
> AND
>
> (concept OR "case studies" OR "case study" OR "definition" . . .)

Once we completed the searches, the results were downloaded into Zotero folders. The Q1 search generated over 10,000 articles, compared to the other three questions, which each yielded between 300 and 2,000 articles. Our original expectation was that the literature identified through Q1 would be focused on the conceptual underpinnings of CI, but many results did not discuss features of CI or were opinion or advocacy articles which made them difficult to synthesis. We concluded that the more specific search terms for the remaining questions helped yield more relevant studies for our analysis, so we moved forward using only the Q2, Q3, and Q4 search results, which were intended to contain empirical studies. Although we did not code the studies resulting from Q1, we drew on some of ideas we had intended to capture in Q1 in developing the coding schemes for the articles we ultimately included in the

synthesis. After deduplication of the Q2, Q3, and Q4 search results, 2,578 search results remained for screening. The results were then downloaded into a Microsoft Access database, which was used for all screening and the first round of coding.

Phase 2. Screened Study Abstracts for Relevance and Performed Initial Round of Full-Text Coding

We scanned the search results' title, source title, abstract text, and date, categorizing each source as related to education (prekindergarten through postsecondary), healthcare, or both, eliminating patently irrelevant search results. For example, we excluded Miller et al. (2015)—"Effectiveness of Interventions for Convicted DUI Offenders in Reducing Recidivism: A Systematic Review of the Peer-Reviewed Scientific Literature"—and Youn et al. (2019)—"Practical Solutions for Sustaining Long-Term Academic-Community Partnerships."

After this initial screening process, 1,731 studies remained. Because of the large volume of studies remaining, we decided to eliminate the sources that discussed CI efforts in healthcare only. To determine whether the full-text article should be coded, we screened the remaining study abstracts using the following two inclusion criteria:

- Aligns with our working definition of CI, continuous (quality) improvement, improvement science, or implementation science
- Aligns with at least one guiding question

Our working definitions were as follows:

> **Best and Dunlap (2014, p. 1)**: In education, continuous improvement can refer to a school, district, or other organization's ongoing commitment to quality improvement efforts that are evidence-based, integrated into the daily work of individuals, contextualized within a system, and iterative (Park et al., 2013).
>
> **Park et al. (2013, p. 5)**: [Continuous] (quality) improvement is the act of integrating quality improvement into the daily work of individuals in the system. It is a characteristic, or rather a set of three characteristics, of an organization that is both designed and managed to improve over time vis-à-vis desired outcomes in light of a specific system aim. "Continuous", in this sense, is a qualifying adjective of quality improvement work which connotes three organizational characteristics: 1) the frequency of quality improvement work; 2) the depth and extent of its integration at different levels of the organization; and 3) the extent of contextualization within a system of work processes.

Regional Educational Laboratory Program West (2017): Improvement science is a problem-solving approach centered on continuous inquiry and learning. Change ideas are tested in rapid cycles, resulting in efficient and useful feedback to inform system improvements.

Fogarty International Center (2021): Implementation science is the study of methods to promote the adoption and integration of evidence-based practices, interventions and policies into routine health care and public health settings. Implementation research plays an important role in identifying barriers to, and enablers of, effective global health programming and policymaking, and leveraging that knowledge to develop evidence-based innovations in effective delivery approaches.

Once we eliminated the health-care-related studies and completed this phase of screening, 649 sources remained. We then furthered narrowed the pool by restricting the review to studies that

- focused on CI, continuous (quality) improvement, or improvement science and
- were based on empirical research (not purely theoretical).

Fifty-seven studies met the criteria. We excluded the following types of studies: theoretical studies, implementation guides, studies that focused on implementation science (e.g., a study that examined strategies to promote or measure the implementation of an innovation), and literature reviews. Thus, we included empirical studies that focused on CI, continuous quality improvement, or improvement science. When it was unclear if a study should be included or excluded, it was sent to an adjudicator who would make a final determination.

Phase 3. Screened Full-Text Studies and Performed Second Round of Coding

We coded the remaining 57 studies in a Microsoft Access database. The codes used for the initial round of coding are in Appendix D. During this phase, we exported the resulting categorical and thematic data from the Access database to a Microsoft Excel file. Reviewers split this file into separate Excel files based on data that corresponded to their assigned guiding questions; for example, data related to CI team membership were exported into one data file. Each reviewer then analyzed those data using Excel and NVivo to develop common categories and themes that corresponded to their assigned guiding questions. We examined the raw data to explore any potentially missing codes, inconsistencies, or apparent discrepancies.

The coding of the 57 studies revealed that some studies included in the pool were not actually studies of CI efforts as intended. This led us to conduct one final round of screening, drawing on the initial coding. In this round, we restricted the pool of studies to those that met the following additional criteria:

- Focused on CI initiatives carried out by practitioners (e.g., teachers, school-level administrators)
- Discussed practitioners implementing one or more inquiry cycles (e.g., PDSA cycles)

We excluded studies in which the CI effort was conducted by external researchers who did not identify as school, district, or state education practitioners as well as studies in which practitioners did not test specific interventions or change ideas using inquiry cycles. For example, Page and Eadie (2019) examined the features of two collaborative interdisciplinary research projects that supported early childhood educators engaged in the process of continuous improvements; however, the early childhood educators did not engage in inquiry cycles as part of either project. If it was unclear if the studies should be included or excluded, they were sent through the established adjudication process. After these rounds of screening, 34 studies remained for synthesis.

Synthesis

Our findings are based on the 34 empirical studies that satisfied our screening criteria. The codes used to structure the two rounds of coding are in Appendix D. We synthesized the literature by developing categories and themes using two data coding techniques: a priori coding (i.e., categorizing text using terms developed beforehand) and emergent coding (i.e., reading through text to identify emerging categories or themes).

The a priori codes were developed to identify the key features of each study (e.g., setting of CI effort), whereas others were developed to align with the Bill & Melinda Gates Foundation's terminology and Networks for School Improvement strategy (e.g., change ideas or interventions used to implement CI). We employed an emergent coding approach to help us identify themes, patterns, and categories related to our broader guiding questions and concepts (e.g., conditions that support or constrain CI implementation).

The counts in the text and exhibits are based on the number of studies that report specific features or characteristics of CI. As described in the Introduction and Appendix C, 25 of the 34 studies we reviewed described distinct CI initiatives. The other nine studies described three initiatives: the CORE-PACE research partnership (five studies), the National Center on Scaling Up Effective Schools (two studies), and the Quest Project (two studies).

Limitations

Our literature searches were limited to databases available via the EBSCO interface. We did not limit our search to articles that had their full text available in EBSCO; we also searched titles, abstracts, and subject keywords to include articles that may have been relevant. When full-text articles were not available given our institution's subscriptions, we requested these articles from other libraries through interlibrary loan. However, full-text availability also limited the ability to fully screen and code all the items.

The 34 empirical studies included in the review vary in rigor. Thirty-one of the 34 studies were case studies and qualitative research that relied mostly on interviews, surveys, research observations, and artifact analysis. The other three studies explicitly employed comparison groups (Arbour et al., 2016; Brown, 2013; Gallagher et al., 2019).

In addition, the coding of each article was based on the judgment of individual coders. To ensure reliability, we provided definitions accompanying each code. However, researchers' discretion in determining when to count a study within one category versus another may have created some inconsistencies.

Finally, if a study did not report on a specific analysis topic (e.g., characteristics of students who were the focus of the CI effort), we could not categorize the study for that particular review section.

Appendix B. Studies Included in the Review

Exhibit B1. Studies Selected for Second Phase Coding and Selected Codes

Citation	Setting	Length of CI initiative	Focus of CI effort	Organizational level of CI effort	Focal grade(s) of CI effort
Adams, D. (2013). *Creating a new learning community: A case study of school leaders' perceptions of the Baldrige Criteria for Performance Excellence and their implementation in a suburban school district* [Unpublished doctoral dissertation]. Northern Illinois University.	Suburban	15 years	• Improve students' academic skills (English language arts [ELA]/reading)	School and district	PK–Grade 8
Aggarwal, A. K., & Lynn, S. A. (2012). Using continuous improvement to enhance an online course. *Decision Sciences Journal of Innovative Education*, *10*(1), 25–48.	Urban	1 year	• Improve students' academic skills	Postsecondary	n/a
Aguilar, J., Nayfack, M., & Bush-Mecenas, S. (2017). *Exploring improvement science in education: Promoting college access in Fresno Unified School District*. Stanford Graduate School of Education, Policy Analysis for California Education.	Urban	1 year	• Increase students' access to postsecondary opportunities	District	Grade 12
Arbour, M., Yoshikawa, H., Atwood, S., Duran Mellado, F. R., Godoy Ossa, F., Trevino Villareal, E., & Snow, C. E. (2016). *Improving quality and child outcomes in early childhood education by redefining the role afforded to teachers in professional development: A continuous quality improvement learning collaborative among public preschools in Chile*. Society for Research on Educational Effectiveness.	Rural and urban	2 years	• Improve students' academic skills • Improve teacher practice	School	PK

Citation	Setting	Length of CI initiative	Focus of CI effort	Organizational level of CI effort	Focal grade(s) of CI effort
Arnold, A. J. (2010). *Exploring the use of the eight-step process in the area of third grade communication arts* [Unpublished doctoral dissertation]. Lindenwood University.	Suburban	1 year	• Improve students' academic skills (ELA/reading) • Improve teacher practice	School	Grade 3
Aumiller, B. E. (2008*). Implementation of the Baldrige education framework for performance excellence: A case study of one Midwestern school district leadership team's continuous improvement efforts* (Publication No. 3314727) [Doctoral dissertation, University of Illinois at Urbana-Champaign]. ProQuest.	Suburban	4 years	• Other	School and district	K–Grade 8
Brown, J. V. (2013). The continuous improvement model: A K–12 literacy focus [Unpublished doctoral dissertation]. Lipscomb University.	Urban	1 year	• Improve students' academic skills (ELA/reading) • Improve teacher practice	School	Grades 4–10
Cannata, M., Redding, C., & Rubin, M. (2016). *Continuous improvement in action: Educators' evidence use for school improvement*. Society for Research on Educational Effectiveness.	Not stated	3 years	• Improve students' social-emotional or behavioral skills	School	Grades 9–12
Edwards, A. R., Sandoval, C., & McNamara, H. (2015). Designing for improvement in professional development for community college developmental mathematics faculty. *Journal of Teacher Education*, *66*(5), 466–481.	Not stated	1 year	• Other	Postsecondary	n/a

Citation	Setting	Length of CI initiative	Focus of CI effort	Organizational level of CI effort	Focal grade(s) of CI effort
Ell, F., & Meissel, K. (2011). Working collaboratively to improve the learning and teaching of mathematics in a rural New Zealand community. *Mathematics Education Research Journal*, *23*(2), 169–187.	Rural	1 year	• Improve students' academic skills (math) • Improve teacher practice	School	K–Grade 8
Ermeling, B. A. (2010). Tracing the effects of teacher inquiry on classroom practice. *Teaching and Teacher Education*, *26*(3), 377–388.	Urban	1 year	• Improve teacher practice	School	Grades 9–12
Gallagher, H. A., & Cottingham, B. W. (2019). *Learning and practicing continuous improvement: Lessons from the CORE districts*. Stanford Graduate School of Education, Policy Analysis for California Education.	Urban	4 years	• Improve students' academic skills (ELA/reading, math)	School and district	Grades 4–8
Gallagher, H. A., Gong, A., Hough, H. J., Kennedy, K., Allbright, T., & Daramola, E. J. (2019). *Engaging district and school leaders in continuous improvement: Lessons from the 2nd year of implementing the CORE Improvement Community*. Stanford Graduate School of Education, Policy Analysis for California Education.	Urban	2 years	• Improve students' academic skills (math)	School and district	Grades 4–8
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.	Urban	3 years	• Other	School	PK–Grade 12
Hardy, I. (2016). In support of teachers' learning: Specifying and contextualising teacher inquiry as professional practice. *Asia-Pacific Journal of Teacher Education*, *44*(1), 4–19.	Urban	1 year	• Improve teacher practice	School	Grade 5

Citation	Setting	Length of CI initiative	Focus of CI effort	Organizational level of CI effort	Focal grade(s) of CI effort
Harrison, C., Wachen, J., Brown, S., & Cohen-Vogel, L. (2019). A view from within: Lessons learned from partnering for continuous improvement. *Teachers College Record*, *121*(9), 1–38.	Urban	5 years	• Improve students' academic skills • Improve students' social-emotional or behavioral skills	School and district	Grades 9–12
Howley-Rowe, C. (2000a). *Bending Knee Elementary: A case study of the Quest Network*. AEL.	Rural	3 years	• Improve students' academic skills (ELA/reading) • Improve teacher practice	School	PK–Grade 6
Howley-Rowe, C. (2000b). *Bowman Elementary: A case study of the Quest Network*. AEL.	Urban	3 years	• Improve students' academic skills (ELA/reading) • Improve teacher practice	School	K–Grade 5
Huffman, D., & Kalnin, J. (2003). Collaborative inquiry to make data-based decisions in schools. *Teaching and Teacher Education*, *19*(6), 569–580.	Not stated	1 year	• Improve students' academic skills (mathematics) • Improve teacher practice	School and district	K–Grade 12
Kaufman, E. K., Cash, C., Coartney, J., Ripley, D., Guy, T., Glenn, W., Mitra, S., & Anderson, J. C., II. (2019). Planning to create a culture of continuous improvement with the Department of Defense Education Activity. *Educational Planning*, *26*(4), 5–19.	Not stated	5 years	• Improve students' academic skills	Federal government	K–Grade 12

Citation	Setting	Length of CI initiative	Focus of CI effort	Organizational level of CI effort	Focal grade(s) of CI effort
Kennedy, K., & Gallagher, H. A. (2019). *Leadership that supports continuous improvement: The case of Ayer Elementary*. Stanford Graduate School of Education, Policy Analysis for California Education.	Urban	1 year	• Improve students' academic skills (mathematics) • Improve students' social-emotional or behavioral skills	School	K–Grade 6
Liang, S., Mira, R. V., Prasad, P. V., & Patterson, C. L. (2019). Improving our practice as mathematics teacher educators through teaching research. *International Journal for the Scholarship of Teaching and Learning*, *13*(2), Article 12.	Not stated	2 years	• Improve teacher practice	Postsecondary	n/a
Park, S., Hironaka, S., Carver, P., & Nordstrum, L. (2013). *Continuous improvement in education*. Carnegie Foundation for the Advancement of Teaching.	Urban	2 years	• Improve students' academic skills • Improve students' social-emotional or behavioral skills	School and district	Grades 9–12
Park, V. (2019). *Bridging the knowing-doing gap for continuous improvement: The case of Long Beach Unified School District*. Stanford Graduate School of Education, Policy Analysis for California Education.	Urban	1 year	• Improve students' academic skills (mathematics) • Improve teacher practice	School and district	PK–Grade 12
Ray, J. L. (2014). *A multiple case study of a Baldrige-based peer review process in an urban school district* (Publication No. 3536718) [Doctoral dissertation, Sam Houston State University]. ProQuest.	Urban	6 years	• Improve students' academic skills	School and district	PK–Grade 12

Citation	Setting	Length of CI initiative	Focus of CI effort	Organizational level of CI effort	Focal grade(s) of CI effort
Redding, C., & Viano, S. (2018). Co-creating school innovations: Should self-determination be a component of school improvement? *Teachers College Record, 120*(11), 1–32.	Urban	2 years	• Improve students' academic skills • Improve students' social-emotional or behavioral skills • Increase students' access to postsecondary opportunities	School and district	Grades 9–12
Rohanna, K. (2017). Breaking the "adopt, attack, abandon" cycle: A case for improvement science in K–12 education. *New Directions for Evaluation*, *153*, 65–77.	Urban	1 year	• Improve students' academic skills	School and district	Grades 6–8
Rohanna, K. (2018). *Preparing schools to successfully participate in networked improvement communities: A case study of year 1 of a math instructional network.* University of California–Los Angeles.	Urban	1 year	• Improve students' academic skills (mathematics)	School	Grades 6–9
Shearer, K. (2018). *Creating systems of continuous improvement: A phenomenological study of California districts engaged in transforming into learning organizations* [Unpublished doctoral dissertation]. Brandman University.	Rural, suburban, and urban	2 years	• Other	School and district	PK–Grade 12
Sinapuelas, M. L., Lardy, C., Korb, M. A., Bae, C. L., & DiStefano, R. (2019). Developing a three-dimensional view of science teaching: A tool to support preservice teacher discourse. *Journal of Science Teacher Education*, *30*(2), 101–121.	Not stated	1 year	• Improve teacher practice	Postsecondary	n/a

Citation	Setting	Length of CI initiative	Focus of CI effort	Organizational level of CI effort	Focal grade(s) of CI effort
Taylor, R. T., & Gordon, W. R., II. (2014). 4 schools, 1 goal: University-district partnership nets results for struggling readers. *The Learning Professional*, *35*(3), 16–20.	Urban	2 years	• Improve students' academic skills (ELA/reading) • Improve teacher practice	School	Grades 9–10
Wallenstein, J. (2018*). Investing in school learning: The New York City Department of Education's Learning Partners Program*. Columbia University.	Urban	3 years	• Improve students' academic skills (ELA/reading) • Improve students' social-emotional or behavioral skills • Improve teacher practice • Increase students' access to postsecondary opportunities • Improve parent and family engagement	School	PK–Grade 12
Walters, M. B. (2012). *The impact of the data teams process on student mathematics achievement* (Publication No. 3575036) [Doctoral dissertation, Texas Southern University]. ProQuest.	Not stated	3 years	• Improve students' academic skills (math)	School	Grade 12

Citation	Setting	Length of CI initiative	Focus of CI effort	Organizational level of CI effort	Focal grade(s) of CI effort
Wilcox, K. C., Lawson, H. A., & Angelis, J. I. (2017). COMPASS-AIM: A university/P–12 partnership innovation for continuous improvement. *Peabody Journal of Education*, *92*(5), 649–674. https://doi.org/10.1080/0161956X.2017.1368654	Rural, suburban, and urban	5 years	• Improve students' academic skills (ELA/reading, mathematics, science) • Improve students' social-emotional or behavioral skills • Improve teacher practice • Improve parent and family engagement	School and district	PK–Grade 12

Appendix C. Three Continuous Improvement Initiatives Examined in Nine Studies

Exhibit C1. Three Continuous Improvement Initiatives Were Discussed in Nine Studies

Sources	CI initiative name	Initiative description
Aguilar et al. (2017); Gallagher & Cottingham (2019); Gallagher et al. (2019); Kennedy & Gallagher (2019); Park (2019)	CORE-PACE Research Partnership	In October 2015, Policy Analysis for California Education (PACE) launched a research partnership with the CORE districts in California. The CORE-PACE Research Partnership is focused on producing research that informs CI in the CORE districts and policy and practice in California and beyond.
Harrison et al. (2019); Tichnor-Wagner et al. (2017)	National Center on Scaling Up Effective Schools (NCSU)	In a 5-year period, NCSU worked to enact an improvement process in two large urban districts in the United States. In doing so, NCSU worked shoulder-to-shoulder with practitioners in each district, teams of developers from a large international development center, and researchers from several universities. Together, they developed research-based educational innovations, adapted and implemented them in several school contexts, and expanded them to new school contexts across time.
Howley-Rowe (2000a, 2000b)	Quest Project	As part of its Regional Educational Laboratory (REL) contract to develop a framework for continuous school improvement in its four-state region, AEL Inc. staff designed the Quest Project. Based on principles of inquiry, collaboration, and action research, Quest proposed to support and investigate ongoing school improvement efforts through twice-yearly conferences (which staff renamed rallies), summer symposia, a scholars program, visits to participating schools, communication via electronic mailing list and mailings, and the creation of a Quest network of schools.

Appendix D. Codes for First and Second Phases of Coding

First Phase

Exhibit D1. Codes Used for the First Phase of Coding

Field description	Coding options
1. Does the study focus on continuous improvement (CI)/improvement science?	0 = No; 1 = Yes
1a. If no, please describe why this study is still relevant to review.	1 = Still relevant; 2 = Exclude
1b. If the study is "still relevant," explain.	Open-ended response
2. Does the study define CI/improvement science/implementation science?	0 = No; 1 = Yes
2a. If yes, what is the definition?	Open-ended response
3. Is the study a research study?	0 = No; 1 = Yes
3a. If no, please describe the context and say why this study is still relevant to the review.	1 = Study context; 2 = Exclude
3b. If the study has a "study context," explain.	Open-ended response
4. What are the research questions for the study?	Open-ended response
5. What was the design of the study?	1 = Randomized controlled trial 2 = Quasi-experimental design 3 = Correlational 4 = Case study 5 = Literature review/ systematic review/meta-analysis 6 = Mixed methods 7 = Other experimental 8 = Other nonexperimental 9 = Qualitative study
5a. Other experimental (describe)	Open-ended response
5b. Other nonexperimental (describe)	Open-ended response
6. Describe the context and time period for the study.	Open-ended response
7a. Age group or grade band: Ages 0–3	1 = True
7b. Age group or grade band: Prekindergarten	1 = True
7c. Age group or grade band: Kindergarten	1 = True
7d. Age group or grade band: Grades 1–2	1 = True

Field description	Coding options
7e. Age group or grade band: Grades 3–5	1 = True
7f. Age group or grade band: Middle school (Grades 6–8)	1 = True
7g. Age group or grade band: High school (Grades 9–12)	1 = True
7h. Age group or grade band: Postsecondary	1 = True
7i. Age group or grade band: Adults (e.g., teachers, administrators)	1 = True
8. If students were in the study sample, describe.	Open-ended response
8a. Specific student demographics for study sample: students with disabilities	1 = True
8b. Specific student demographics for study sample: Black/African American	1 = True
8c. Specific student demographics for study sample: Latinx	1 = True
8d. Specific student demographics for study sample: Asian/Pacific Islander	1 = True
8e. Specific student demographics for study sample: Native/indigenous populations	1 = True
8f. Specific student demographics for study sample: White	1 = True
8g. Specific student demographics for study sample: Low income	1 = True
8h. Specific student demographics for study sample: Homeless	1 = True
8i. Specific student demographics for study sample: Immigrant/refugee	1 = True
8j. Specific student demographics for study sample: Other	1 = True
8k. Specific student demographics for study sample: Other (specify)	Open-ended response
9. If teacher/adult is in the study sample, describe.	Open-ended response
10a. Which, if any, of the following outcomes were measured? Student academic (e.g., test scores, grades)?	1 = True
10b. Which, if any, of the following outcomes were measured? Student social-emotional?	1 = True
10c. Which, if any, of the following outcomes were measured? Graduation rates?	1 = True
10d. Which, if any, of the following outcomes were measured? Attendance?	1 = True
10e. Which, if any, of the following outcomes were measured? Grade promotion?	1 = True

Field description	Coding options
10f. Which, if any, of the following outcomes were measured? College-related (e.g., enrollment, application, FAFSA [Free Application for Federal Student Aid] completion)?	1 = True
10g. Which, if any, of the following outcomes were measured? Other student behavior?	1 = True
10h. Which, if any, of the following outcomes were measured? Teacher behavior?	1 = True
10i. Which, if any, of the following outcomes were measured? Other?	1 = True
10j. Which, if any, of the following outcomes were measured? Other (specify)?	Open-ended response
11a. Which types of data were collected? Collected by CI team: student administrative data (e.g., test scores, attendance data)?	1 = True
11b. Which types of data were collected? Collected by CI team: student surveys?	1 = True
11c. Which types of data were collected? Collected by CI team: teacher surveys?	1 = True
11d. Which types of data were collected? Collected by CI team: teacher observation?	1 = True
11e. Which types of data were collected? Collected by CI team: student interviews/focus groups?	1 = True
11f. Which types of data were collected? Collected by CI team: teacher/adult interviews or focus groups?	1 = True
11g. Which types of data were collected? Collected by CI team: artifact analysis?	1 = True
11h. Which types of data were collected? Collected by CI team: researcher-led observations?	1 = True
11i. Which types of data were collected? Collected by CI team: other?	1 = True
11j. Which types of data were collected? Collected by CI team: other (specify)?	Open-ended response
12a. Which types of data were collected? Collected by researchers: student administrative data (e.g., test scores, attendance data)?	1 = True
12b. Which types of data were collected? Collected by researchers: student surveys?	1 = True
12c. Which types of data were collected? Collected by researchers: teacher surveys?	1 = True

Field description	Coding options
12d. Which types of data were collected? Collected by researchers: teacher observation?	1 = True
12e. Which types of data were collected? Collected by researchers: student interviews/focus groups?	1 = True
12f. Which types of data were collected? Collected by researchers: teacher/adult interviews or focus groups?	1 = True
12g. Which types of data were collected? Collected by researchers: artifact analysis?	1 = True
12h. Which types of data were collected? Collected by researchers: researcher-led observations?	1 = True
12i. Which types of data were collected? Collected by researchers: other?	1 = True
13. Share details about data collected by CI team and/or researchers.	Open-ended response
14. If specific measures were developed or used, describe.	Open-ended response
15a. What were they trying to improve?	Open-ended response
15b. What was the practice and/or intervention they were trying to test?	Open-ended response
15c. Did the effort involve implementing a preexisting intervention (e.g., a branded intervention, curriculum, or instructional approach)?	0 = No; 1 = Yes
15d. If yes, name.	Open-ended response
16. At what organizational level did CI take place? (i.e., Was there a team at the district level, one at each school, one at each grade level in a school, or other?)	Open-ended response
17. Does the study describe the CI team?	0 = No; 1 = Yes
17a. Who were the members? Teachers?	1 = True
17b. Who were the members? Administrators?	1 = True
17c. Who were the members? Other staff?	1 = True
17d. Who were the members? District personnel?	1 = True
17e. Who were the members? Outside CI coaches?	1 = True
17f. Who were the members? Other partners?	1 = True
17g. Who were the members? Community members?	1 = True
17h. Who were the members? Parents?	1 = True

Field description	Coding options
17i. Who were the members? Students?	1 = True
17j. Who were the members? Other?	1 = True
17k. Who were the members? Other (specify)?	Open-ended response
18. Does the study describe how the CI team members were chosen?	0 = No; 1 = Yes
18a. If yes, how?	Open-ended response
19a. Who were the team leaders? Teachers?	1 = True
19b. Who were the team leaders? Administrators?	1 = True
19c. Who were the team leaders? Outside CI coaches?	1 = True
19e. Who were the team leaders? Other?	1 = True
19f. Who were the team leaders? Describe any "other" choices.	Open-ended response
20. Any notes about the CI team?	Open-ended response
21a. Does the study examine the processes used to implement CI? Planning time?	1 = True
21b. Does the study examine the processes used to implement CI? Knowledge of CI process?	1 = True
21c. Does the study examine the processes used to implement CI? Knowledge of staff capacity?	1 = True
21d. Does the study examine the processes used to implement CI? Coaching?	1 = True
21e. Does the study examine the processes used to implement CI? Rubrics?	1 = True
21f. Does the study examine the processes used to implement CI? Handbooks?	1 = True
21g. Does the study examine the processes used to implement CI? Workshops on CI?	1 = True
21h. Does the study examine the processes used to implement CI? Other?	1 = True
21i. Does the study examine the processes used to implement CI? Other (specify)?	Open-ended response
21j. Share details about processes used to implement CI.	Open-ended response
22. Does the study demonstrate use of the Bill & Melinda Gates Foundation's Six Core Parameters of CI? An understanding of the problem?	0 = No; 1 = Yes

Field description	Coding options
22a. If yes, explain.	Open-ended response
23. Does the study demonstrate use of the Bill & Melinda Gates Foundation's Six Core Parameters of CI? A clear and specific aim centered on achieving equitable outcomes for students of color and/or students from low-income backgrounds?	0 = No; 1 = Yes
23a. If yes, explain.	Open-ended response
24. Does the study demonstrate use of the Bill & Melinda Gates Foundation's Six Core Parameters of CI? Equity-centered theory of practice improvement?	0 = No; 1 = Yes
24a. If yes, explain.	Open-ended response
25. Does the study demonstrate use of the Bill & Melinda Gates Foundation's Six Core Parameters of CI? Disciplined inquiry cycles?	0 = No; 1 = Yes
25a. If yes, explain.	Open-ended response
26. Does the study demonstrate use of the Bill & Melinda Gates Foundation's Six Core Parameters of CI? Collaborative and diverse teams?	0 = No; 1 = Yes
26a. If yes, explain.	Open-ended response
27. Does the study demonstrate use of the Bill & Melinda Gates Foundation's Six Core Parameters of CI? Use of data from multiple sources?	0 = No; 1 = Yes
27a. If yes, explain.	Open-ended response
28. If (4) Disciplined inquiry cycles is yes, does the study report on inquiry cycles?	0 = No; 1 = Yes
28a. If yes, how many per year?	Open-ended response
28b. If yes, how long were they?	Open-ended response
29. Does the study examine whether there was improvement in specific outcomes?	0 = No; 1 = Yes
29a. If yes, did the outcomes show improvement?	0 = No; 1 = Yes; 2 = Mixed results
29b. Share details about outcomes that show improvement.	Open-ended response
30a. Did the study examine whether particular supports and resources improved outcomes? If so, which supports and resources were key in facilitating outcomes? Planning time?	1 = True
30b. Did the study examine whether particular supports and resources improved outcomes? If so, which supports and resources were key in facilitating outcomes? Knowledge of CI process?	1 = True

Field description	Coding options
30c. Did the study examine whether particular supports and resources improved outcomes? If so, which supports and resources were key in facilitating outcomes? Knowledge of staff capacity?	1 = True
30d. Did the study examine whether particular supports and resources improved outcomes? If so, which supports and resources were key in facilitating outcomes? Coaching?	1 = True
30e. Did the study examine whether particular supports and resources improved outcomes? If so, which supports and resources were key in facilitating outcomes? Rubrics?	1 = True
30f. Did the study examine whether particular supports and resources improved outcomes? If so, which supports and resources were key in facilitating outcomes? Handbooks?	1 = True
30g. Did the study examine whether particular supports and resources improved outcomes? If so, which supports and resources were key in facilitating outcomes? Workshops on CI?	1 = True
30h. Did the study examine whether particular supports and resources improved outcomes? If so, which supports and resources were key in facilitating outcomes? Other?	1 = True
30i. Did the study examine whether particular supports and resources improved outcomes? If so, which supports and resources were key in facilitating outcomes? Other (specify)?	Open-ended response
30j. Share details about particular supports and resources that improved outcomes.	Open-ended response
31. Were there additional findings outside of what they intended to measure?	Open-ended response
32. Does the study discuss the degree to which CI was implemented?	0 = No; 1 = Yes
32a. If yes, how well was it implemented?	1 = Well; 2 = Not well; 3 = Mixed results
33a. What were the barriers to implementation? Time?	1 = True
33b. What were the barriers to implementation? Funds?	1 = True
33c. What were the barriers to implementation? Staff capacity?	1 = True
33d. What were the barriers to implementation? Other?	1 = True
33e. What were the barriers to implementation? Other (specify)?	Open-ended response
33f. More details about barriers to implementation?	Open-ended response

Field description	Coding options
34. Does the study discuss diffusion of CI interventions/practices beyond the CI team?	0 = No; 1 = Yes
34a. To others in the same building?	0 = No; 1 = Yes
34a1. If yes, describe the process.	Open-ended response
34b. To others beyond the school building?	0 = No; 1 = Yes
34b1. If yes, describe the process.	Open-ended response
34c. Did they study how effective this diffusion was?	0 = No; 1 = Yes
34c1. If yes, describe the results of the diffusion.	Open-ended response
35. What were the lessons learned for adoption in other places?	Open-ended response
36. Reviewer comments	Open-ended response

Second Phase

Exhibit D2. Codes Used for the Second Phase of Coding

Code Name	Coding options
1. Definition of CI used in study (if stated)	Open code
2. Name of CI initiative (if stated)	Open code
3. Time period of CI effort	Open code
4. Evidence of author involvement in CI effort	a. Author organization developed CI framework or initiative b. Author provided guidance and/or expertise to practitioners c. No evidence
5. Setting of CI effort	a. Not stated b. Rural c. Suburban d. Urban
6. Organizational level of CI effort	a. School level b. District level c. Postsecondary level d. Local government e. Federal government

Code Name	Coding options
7. Grade band of students who were the focus of CI effort	a. Prekindergarten b. Elementary school (Grades K–5) c. Middle school (Grades 6–8) d. High school (Grades 9–12)
8. Demographics or characteristics of students who were focus of CI effort	a. Asian/Pacific Islander students b. Black or African American students c. Latino/a/x students d. Native/indigenous populations e. English learners f. Students experiencing homelessness g. Students from low-income households h. Students in special education i. Students from low-performing schools j. Students from high-performing schools k. Other—open code l. N/A—No clear focal student population
9. Focus of CI effort	a. Improve students' academic skills b. Improve students' social-emotional or behavioral skills c. Improve teacher instruction d. Increase students' access to postsecondary opportunities e. Improve parent and family engagement
10. Change ideas or interventions used to implement CI	a. Collaboration routines—Level 1 i. Create high-functioning teams—Level 2 ii. Staff development and support—Level 2 iii. Structures-routines for collaborative work—Level 2 b. Supportive environments—Level 1 i. College access and affordability—Level 2 ii. Create a college-going culture—Level 2 iii. Promote health, safety, and well-being—Level 2 iv. Student agency, identity, and mindset—Level 2

Code Name	Coding options
11. Members of CI team	a. CI coaches or facilitators b. District personnel c. Guidance counselors d. Instructional coaches e. Researchers f. School-level administrators g. Teachers h. University faculty i. University faculty administrators j. Other—open code
12. Data collected by authors to measure CI effort	a. Artifact analysis b. Practitioner interview data c. Practitioner observation d. Practitioner survey data e. Student academic data f. Student activity data g. Student interview data h. Student survey data
13. Data collected as part of CI effort by CI team members	a. Artifact analysis b. Practitioner interview data c. Practitioner observation d. Practitioner survey data e. Student academic data f. Student activity data g. Student interview data h. Student survey data
14. Processes used to implement CI effort	a. Coaching b. Collaborative learning opportunities c. Peer observation d. Professional development e. Use of tools
15. Number of inquiry cycles conducted by practitioners	Open code

Code Name	Coding options
16. Evidence of diffusion	a. Within the school/district/university engaged in the CI effort b. With other schools/districts/universities engaged in the CI effort c. With other schools/districts/universities not involved in the CI effort
17. Conditions that support or constrain CI implementation	a. Conditions that constrained CI implementation i. Lack of fiscal resources ii. Lack of staff buy-in iii. Lack of staff capacity or expertise iv. Time constraints b. Conditions that supported CI implementation i. Collaborative teachers 1. Quality professional development 2. Opportunities for collaboration and reflection on practice 3. Respectful and trusting professional relationships 4. Access to new ideas ii. Effective leaders 1. Supportive leadership 2. Shared leadership 3. Organizational coherence
18. Conclusions on whether the CI effort led to improved outcomes	a. Improved the organizational capacity to implement CI b. Improved student academic outcomes c. Improved teacher instruction d. Improved family and parent engagement e. Increased students' access to postsecondary opportunities f. Other

About the American Institutes for Research

Established in 1946, with headquarters in Arlington, Virginia, the American Institutes for Research® (AIR®) is a nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance to solve some of the most urgent challenges in the U.S. and around the world. We advance evidence in the areas of education, health, the workforce, human services, and international development to create a better, more equitable world. The AIR family of organizations now includes IMPAQ, Maher & Maher, and Kimetrica. For more information, visit AIR.ORG.

American Institutes for Research®
1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
202.403.5000 | **AIR.ORG**