AMERICAN INSTITUTES FOR RESEARCH®

Evidence for Networked Improvement Communities

A Systematic Review of the Literature

JANUARY 2020

Amy Feygin | Liz Nolan | Alexandra Hickling | Lawrence Friedman

MAKING RESEARCH **RELEVANT**

Evidence for Networked Improvement Communities

A Systematic Review of the Literature

JANUARY 2020

Amy Feygin | Liz Nolan | Alexandra Hickling | Lawrence Friedman

American Institutes for Research®

1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
202.403.5000

www.air.org

Contents

Tables

Introduction

There is growing interest in using improvement science to address complex problems of educational practice. Improvement science is an approach to improvement that involves the systematic testing of changes in practice by using continuous inquiry cycles. Improvement science was first developed to improve industrial manufacturing processes but also has been applied to medicine and engineering (Bryk, 2015). When improvement science is applied in a networked improvement community (NIC)—a type of collaborative research partnership—participants work together to systematically test and refine promising solutions to a common problem of practice in different contexts. Through this process, the participants learn about what works, for whom, and under what conditions, thereby accelerating improvement (Bryk, Gomez, LeMahieu, & Grunow, 2015).

NICs are distinguished by four essential features: (1) NICs are focused on a well-specified common aim; (2) they are guided by a deep understanding of the problem, the system that produces it, and a shared working theory of how to improve it; (3) their work is disciplined by the rigor of improvement science; and (4) they are coordinated to accelerate the development, testing, and refinement of interventions, their rapid diffusion into the field, and their effective integration into varied educational contexts (Bryk et al., 2015). For example, the Carnegie Foundation for the Advancement of Teaching formed the Building a Teacher Effectiveness Network to address the problem of novice teacher development and retention. Leaders from education practice, policy, and research joined forces to identify and address the root causes of low novice teacher retention rates. Through rapid, small-scale tests of change, principals and novice teachers developed a teacher feedback protocol intended to reduce feelings of burnout and increase confidence (Carnegie Foundation for the Advancement of Teaching, n.d.).

The number of NICs has grown considerably over the past few years. The Carnegie Foundation for the Advancement of Teaching has led an effort to educate researchers and practitioners about NICs through annual summits over the past 5 years and the publication of the book *Learning to Improve,* in 2015. Funding from private foundations, such as the Bill & Melinda Gates Foundation and the Nellie Mae Education Foundation, as well federal agencies, such as the National Science Foundation, has further encouraged the growth of NICs. A search of grants awarded by the National Science Foundation's Education and Human Resources Directorate indicates a total of 46 grant abstracts mentioning "networked improvement," with 39 of these in the past 3 years.[1] Apart from these efforts, the requirements of the Every Student Succeeds Act (ESSA) for schools identified as in need of improvement have encouraged schools and districts to consider participation in NICs. Although ESSA does not explicitly mention

[1] The search of grants was conducted in August 2018.

improvement science or NICs, states have flexibility to develop their own strategies for addressing problems of practice. Under these guidelines, schools may choose to adopt and test promising practices through continuous inquiry cycles (Klein, 2018).

However, there have been few efforts to study NICs in education to date and, as a result, little is known about their implementation or outcomes. AIR set out to conduct a systematic review of the empirical research literature on NICs, with the goals of documenting what is known about the implementation and outcomes of NICs and laying a foundation for future research on NICs. In the following sections, we describe our approach to the systematic review, what we found, and the implications of our findings.

Systematic Review Approach

We employed a four-phase process to ensure a thorough and systematic review. During Phase 1, the review team conducted a comprehensive search of the literature to identify studies that carried out empirical research on NICs. During Phase 2, the review team screened study abstracts, using a set of criteria to determine which studies were eligible for further review. During Phase 3, the review team screened the full text of the studies determined to be eligible in Phase 2, using the same criteria. During Phase 4, the review team coded the articles determined to be eligible in Phase 3. Each phase is described in greater detail below.

Phase 1: Literature Search

We began by conducting an extensive search of the literature.[2] We searched the following online education databases to identify articles for inclusion in this systematic review: EBSCO's Education Research Complete, Education Source, and ERIC. To search the databases, we developed the Boolean search string presented in Figure 1, designed to locate studies that described empirical research on NICs.

Figure 1. Boolean Search String for NIC Literature Search

(Networked improvement communit* OR Network improvement communit* OR "networked communit*" OR "networked improvement" OR "Improvement science" OR "improvement communit*" OR "plan do study act" OR "PDSA" OR "inquiry cycle" OR "inquiry cycles") in Select a field, optional
AND
(effect* OR impact* OR evaluat* OR implement* OR success* OR succeed*) in Select a field, optional

This search yielded 318 articles. We tested a less restrictive Boolean search string using "continuous improvement" instead of "networked improvement communities." This search

[2] The search was completed on October 4, 2018, and updated on May 2, 2019.

yielded more than 1,800 articles. Because of our specific interest in NICs, we focused on the restricted set of articles.

Phase 2: Abstract Screening

Trained screeners reviewed abstracts of studies identified in Phase 1 to determine if they met the following criteria for inclusion:

- *Focus:* We included studies in which the primary focus was NICs, improvement science, or a related improvement process, such as Plan-Do-Study-Act (PDSA) cycles or inquiry cycles. We also included studies that focused specifically on education (rather than healthcare, manufacturing, or another field).[3]
- *Type of study:* We included studies that could be characterized as empirical research, defined as "the collection and analysis of either quantitative or qualitative data." Commentaries, literature reviews, theoretical papers, or guides were not eligible for review.
- *Language:* We included only studies written in English.
- *Publication status:* We included both published and unpublished study reports.

Each abstract was screened by two trained reviewers, and discrepancies were resolved by the project director. The reviewers most often excluded abstracts because they were for articles unrelated to NICs. Several terms included in our search string had an alternative but unrelated meaning in education. For example, *inquiry* often referred to "student inquiry"—a common method in science education—rather than "inquiry cycles." Similarly, *networked community* often referred to a "community of learners using a common online learning platform" rather than a "networked improvement community." In addition, many abstracts were excluded because the articles they summarized were not research. Instead, the articles were about the implementation of NICs and, in some cases, lessons learned from their implementation but did not base their findings on the systematic collection and analysis of data. In total, 40 abstracts passed the abstract screening phase.

Phase 3: Full-Text Screening

The review team located full-text articles for the 40 abstracts that met the inclusion criteria. The team then screened the full text articles, using the same inclusion criteria and process as they had for the abstract screening.

It was not always possible to determine on the basis of the abstract whether the study focused on a NIC. In some cases, the term *networked improvement* referred to something unrelated,

[3] Although we limited our search to education databases, we found several articles that were related not to education but to health care or manufacturing processes.

and in other cases, the authors used improvement science within a network of participants but did not explicitly refer to the process as a NIC. Therefore, to assess whether the improvement process described in the study involved a NIC, the review team looked for evidence that the improvement process was designed to have three of the four essential characteristics described by Bryk and colleagues (2015). (The review team did not screen for the second essential characteristic of NICs: that they are guided by a deep understanding of the problem, the system that produces it, and a shared working theory of how to improve it. The articles we reviewed seldom provided sufficient information to determine whether this characteristic was present.)

The first essential characteristic of NICs is that they are focused on a well-specified common aim. To be included in the review, the article needed to describe a community of educators and researchers engaged in addressing a common aim. The third essential characteristic of NICs is that their work is disciplined by the rigor of improvement science. Although we did not require authors to mention improvement science, to be included in the review, the article needed to provide evidence that the community members were engaged in a disciplined improvement process, such as using data to deeply understand a problem and make decisions about how to address the problem. The fourth essential characteristic of NICs is that they are coordinated to accelerate the development, testing, and refinement of interventions, their rapid diffusion into the field, and their effective integration into varied educational contexts. To be included in the review, the article needed to provide evidence of coordination among two or more participants and of participants' meeting on a regular basis to review their progress toward a common aim and to learn from one another's experiences.

In cases in which the study authors did not provide sufficient information to determine whether the criteria had been met, the review team discussed the article and came to a consensus about the way to characterize the study.

After the full-text screening, the team determined that seven articles met the inclusion criteria; these articles moved on to the data extraction and coding phase. Among the 33 excluded articles, more than half failed to meet the criterion that the article had to describe empirical research. As was the case with the abstract screening, studies often described the application of improvement science to a problem of practice and resultant lessons learned from that case but did not systematically collect and analyze data to draw conclusions that could contribute to a broader understanding of improvement science and networked improvement communities.

Phase 4: Data Extraction and Coding

After identifying seven studies that were eligible for inclusion in the review, the review team developed a protocol for coding each study (see Appendix A). The coding protocol included the following components:

- *Study-level information*: Details about the publication, including authors, title, year, and source
- *Network information*: Details about the network, including improvement focus, improvement approach (e.g., PDSA cycles, inquiry cycles), level (e.g., schools, districts, states), number of participants in the network, and setting (e.g., geographic location, urbanicity)
- *Research questions*
- *Type of study:* Implementation or impact study
- *Sample information*: Number of participants (e.g., teachers, schools, or districts) in study and their characteristics
- *Methods information*: Details about data collection, measurement, and the research design
- *Findings*

The same staff members who screened the studies coded them. The final set of coding protocols was analyzed to produce the findings described below.

Findings

Overall, the review team uncovered few studies of the implementation and outcomes of NICs in education. Although the initial search yielded 318 articles, only seven articles met the criteria for inclusion in the review. Three of these articles assessed NIC implementation and four assessed the impact of NICs on student outcomes. Table 1 describes these seven studies.

Table 1. Seven Studies Meeting the Criteria for Inclusion

Study information		NIC description		
Citation	Study type	Year established	Number of schools, districts, or states	Focus of improvement
Hannan, Russell, Takahashi, & Park (2015)	Implementation	Summer 2011	6 schools in a large, urban district that participates in the Building a Teacher Effective Network	Feedback for novice teachers to improve retention
Tichnor-Wagner, Wachen, Cannata, and Cohen-Vogel (2017)	Implementation	2013–14 school year	3 high schools in Broward County (FL) Public Schools, 3 high schools in Fort Worth (TX) Independent School District	Rigorous academic and social-emotional learning for high school students
Rohanna (2017)	Implementation	Not specified	3 middle schools in a California Bay Area school district	Academic perseverance of middle school students
Yamada & Bryk (2016)	Impact	2011–12 school year	19 colleges across 5 states	Developmental math success
Huang & Yamada (2017)	Impact	2011–12 school year	21 colleges across 7 states	Developmental math success
Yamada, Bohannon, Grunow, & Thorn (2018)	Impact	Spring 2012	14 colleges across 8 states	Developmental math success
Ell & Meissel (2011)	Impact	2006	7 schools in a rural community in New Zealand	Quality of instruction in mathematics

Research on NIC Implementation

Three studies assessed the implementation of NICs. All three studies employed a comparative case study design and involved the collection and analysis of data from multiple sources, including interviews, surveys, observations, and PDSA cycle artifacts.

Hannan, Russell, Takahashi, and Park (2015) studied the implementation of the Building a Teaching Effectiveness Network in six focal schools in an urban district. The network focused on improving the retention of beginning teachers by providing supportive structures, such as a formal process for feedback. The researchers wanted to understand the way the schools used improvement science methods to enhance, refine, and integrate the feedback process into their existing system. To do so, they collected in-depth qualitative data from a range of sources: interviews with school principals and school improvement facilitators, artifacts generated from the PDSA cycles, data from an online tracking system, and observations at network meetings. On the basis of their analyses of these data, the researchers grouped schools into three categories: "incomplete enactment" schools had limited engagement with both the feedback process and improvement science; "learning en route to action" schools engaged with both the feedback process and improvement science but struggled to translate their learnings from PDSA cycles into action; and "systematic knowledge into practice" schools were able to integrate the feedback process into their systems, using improvement science methods. The two schools in the systematic-knowledge-into-practice category differed from the other schools in their high-quality documentation of the PDSA process (which facilitated learning across the network) and their ability to act on what they had learned by testing change ideas.

Tichnor-Wagner, Wachen, Cannata, and Cohen-Vogel (2017) studied the implementation of innovation design teams in three high schools in each of two urban districts (Broward County Public Schools and Forth Worth Independent School District). The innovation design teams intended to improve high school student outcomes by increasing access to rigorous academic experiences and social-emotional learning opportunities. The researchers wanted to understand how the will and capacity of innovation design team members to implement PDSA cycles affected the way they engaged in the NIC. The researchers also conducted interviews with district and school practitioners responsible for developing the intervention, researchers, and development specialists after the first year of PDSA implementation. In addition, the researchers conducted observations of PDSA training sessions and collected artifacts, such as field notes, PDSA reflection forms, and feedback forms with Likert-type scales and open-ended questions about participation in the process. The study showed that NIC participants shared a strong belief in the value of PDSA cycles but lacked the will to carry them out. Participants found the process to be laborious, discouraging, and redundant. Some felt insulted that PDSA cycles were being taught to them as something new, even though these participants had been engaging in similar continuous improvement processes for some time. Others resisted the process because they worried that data being collected for the "study" part of the PDSA cycle would be used for accountability purposes.

Finally, Rohanna (2017) studied a California Bay Area school district's efforts to increase the academic perseverance of middle school students. The NIC comprised principals and counselors from five middle schools and central office administrators. To understand the factors that facilitated or constrained the implementation of PDSA cycles, the author collected data from principal interviews, participant observations, and short surveys. She found several themes across the schools, including limited time for teacher collaboration, cultural barriers, limited time among principals, and the importance of an implementation manager. The principal who was most successful at implementing PDSA cycles chose a change idea that built on existing practices at her school and did not require any additional teacher preparation or collaboration. This principal also relied on an assistant principal to serve as her implementation manager. Her sole responsibility in the school was to oversee the implementation of school initiatives, such as the NIC. By shifting responsibility to an implementation manager, the principal was able to overcome the challenge of limited time to manage school initiatives, as well as reduce the burden on teachers.

These studies describe conditions for successful NIC implementation, such as high-quality documentation of PDSA cycles and the use of a dedicated staff member to manage implementation. They also discuss challenges to NIC implementation, such as inconsistent application of PDSA cycles, frustration with an onerous process, and burden on teachers and principals. For future NIC facilitators, the findings provide insights into the way to structure NICs for successful implementation.

Research on NIC Outcomes

Four of the seven studies that met inclusion criteria assessed the outcomes of NICs on students. Three of these four studies focused on two approaches to traditional developmental mathematics courses—Statway® and Quantway®—designed by the Carnegie Foundation for the Advancement of Teaching. Statway is a 1-year college-level set of statistical reasoning courses and Quantway is a 1-semester college-level quantitative reasoning course. Both courses were designed as alternatives to traditional developmental courses for students placed two levels below college-level mathematics. These students were taught using research-based pedagogical approaches that promoted collaborative learning and addressed the social and emotional factors that affected student success in mathematics courses. Faculty who teach Statway and Quantway participate in a comprehensive set of professional development activities and a NIC to continuously improve teaching and learning in the programs.

Yamada and Bryk (2016) compared students who began a traditional developmental mathematics course to students who began Statway 1 year later in 17 colleges. Students were matched using propensity scores within two-level nested models, with students nested within colleges. The researchers found that students who participated in Statway were more likely to earn college-level mathematics credit than their peers who participated in traditional developmental mathematics courses. The positive effects were found across all participating institutions and for all racial/ethnic groups. Huang and Yamada (2017) replicated this analysis after Statway was scaled to 22 additional colleges and universities in its third year of implementation and 26 colleges and universities in its fourth year of implementation. The authors also found that students who participated in Statway were five times more likely to earn college-level mathematics credit than their matched peers, who participated in traditional developmental mathematics courses, and that positive effects were found across all participating institutions and for all racial/ethnic groups.

Also using propensity score matching, Yamada, Bohannon, Grunow, and Thorn (2018) compared students who participated in Quantway with a matched sample of students, who participated in traditional developmental mathematics courses, in 10 colleges. They found that students who participated in Quantway were twice as likely to complete a developmental mathematics course and enroll in a college-level mathematics course the following fall than their matched peers, who participated in traditional developmental mathematics courses. Positive effects were found in all but one college and were found for all racial/ethnic groups.

A fourth article, by Ell and Meissel (2011), assessed an effort to improve mathematics instruction in a rural New Zealand community. Although the authors did not refer to the improvement effort as a NIC, they described a PDSA-like process within a teacher-led cluster of five rural schools. The schools identified a common problem of practice using data, developed a plan for action, and jointly tested the plan through the collection of common data and regular meetings to analyze and interpret the data. Students completed three assessments designed to assess their knowledge of basic facts and place value, which were the targets for the intervention. The authors analyzed all students' change in test scores across the three terms and found that test scores improved. However, no comparison group was constructed to determine what would have happened in the absence of the intervention.

These studies provide very limited evidence on the outcomes of NICs. Although they describe suggestive evidence that the interventions studied led to positive outcomes for students, the methods do not directly test the effect of participation in the NIC on student outcomes. In the three studies focused on Quantway and Statway, the NIC was only one component that factored into the development of the courses and it is not clear that positive associations between participation in Statway and Quantway and attainment of college-level mathematics credit can be attributed to participation in the NIC. The following section delves more deeply into what it would take to design a study that provides evidence on the causal effect of participating in a NIC on student outcomes.

Implications

This review represents a first effort to understand the extant evidence on the implementation and outcomes of NICs.[4] The results of this review raise several implications for education researchers who wish to study NICs.

There is a need for new and better thinking about the best approaches to studying the implementation of NICs. NICs are complex organizations that are difficult to implement. Research that explores variation in the implementation of NICs (e.g., the problems they focus on, their approach to supporting continuous improvement) promises to uncover conditions that promote the successful engagement of educators in systematic improvement efforts. The three studies of implementation identified several conditions and factors that facilitate NIC implementation, as well as conditions that hinder NIC implementation, but more research is needed to have a comprehensive understanding of facilitating and challenging conditions at each stage of the process. One direction work might take is to clarify the conceptual frameworks that guide research on implementation. Two of the three studies of implementation drew on different frameworks. One was based on will and capacity (Tichnor-Wagner et al., 2017) and one on adaptive integration of improvement science methods (Hannan et al., 2015). Another framework described in the literature, developed by Russell and colleagues (2017), posits five critical factors for NIC initiation: (1) learning and using improvement research methods; (2) developing a theory of practice improvement; (3) building a measurement and analytics infrastructure; (4) leading, organizing, and operating the network; and (5) fostering the emergence of culture, norms, and identity consistent with network aims. In designing future implementation studies, research should consider the way the findings from these studies inform existing conceptual frameworks and what modifications may be necessary for greater conceptual clarity.

In addition, the field would benefit from strong thinking on the best research methods for studying NIC implementation. We suggest that NIC facilitators, researchers, and funders hold a series of meetings to discuss and hone methods for studying NIC implementation. We begin

[4] The Center for the Use of Research and Evidence in Education conducted a systematic review of Networked Learning Communities, which were networks of schools working together to address a common problem of practice (Bell, Jopling, Cordingley, Firth, King, & Mitchell, 2006). Although these networks are similar to NICs in their focus on collaboration to accelerate learning, they do not draw on methods from improvement science to engage in systematic tests of strategies for improvement.

this discussion by providing potential implementation questions and data sources for answering them (Table 2). The questions focus on engagement in NIC activities, views of the usefulness and relevance of NIC activities, completion of PDSA cycles, coordination of improvement efforts, and adoption of the practices tested in PDSA cycles. The questions are also are designed to be responsive to the iterative, evolving nature of PDSA cycles. In addition, as with the above implementation studies, we suggest multiple data sources to answer these questions.

Table 2. Example Implementation Questions and Data Sources

Implementation questions	Data source
1. Are NIC members engaged in NIC activities (including NIC meetings and each part of the PDSA cycle)? Does engagement vary over time and across participants (e.g., schools, school districts)?	Participation data Post-PDSA cycle survey
2. Do NIC members find NIC activities (including NIC meetings and each part of the PDSA cycle) to be useful and relevant to their work? Does usefulness and relevance vary over time and across participants (e.g., schools, school districts)?	Post-NIC meeting survey
3. Are NIC members completing PDSA cycles as intended? What challenges are associated with completing PDSA cycles? Does PDSA cycle completion vary over time and across participants (e.g., schools, school districts)?	Artifact review Post-PDSA cycle survey
4. What factors facilitate (or hinder) NIC implementation (e.g., professional development, leadership, resources such as planning time)?	Interviews with network participants
5. Are NIC members coordinating their improvement efforts? What challenges are associated with coordinating improvement efforts? Does coordination vary over time and across participants (e.g., schools, school districts)?	Artifact review Post-PDSA cycle survey
6. Are NIC members adopting the practices targeted by the change ideas tested in PDSA cycles? Does the adoption of practices vary over time and across participants (e.g., schools, school districts)?	Data collection will vary by type of intervention (e.g., an intervention focused on the classroom may use classroom observations or teacher logs)

Innovative methods, such as social network analysis, should be applied to study NIC implementation. NICs are fundamentally social organizations. Although it is possible to engage in improvement without networking, the engagement of educators and organizations across contexts is critical to the type of improvement that NICs aim to generate. Yet, none of the three studies of NIC implementation directly analyzed the social structures of NICs. Researchers should consider using social network analysis to study NIC implementation, which would allow them to answer question about the extent to which participants are interacting and learning from one another.

Social network analysis is a powerful tool for mapping relationships among users (Ansell, Reckhow, & Kelly, 2009; Evans, Rosen, Kesten, & Moore, 2014). Researchers should develop surveys that ask NIC participants, “With whom, if anyone, have you interacted about issues related to this NIC and in what ways, when, where, and why?” Social-network-plotting software then can be used to create visual representations of the network structure, called sociograms. Each node in the network will represents an educator and can be color-coded by his or her organization (e.g., school, school district). This will display the patterns of ties within and among schools, allowing researchers to better understand collaboration among schools.

Researchers also should consider taking advantage of data from online platforms used to facilitate interaction among NIC members. NIC facilitators often use platforms like Slack to encourage NIC participants to share experiences and learn from one another. Researchers should track the number of times each NIC member initiates a new thread or responds to an existing thread, or they should even conduct a social network analysis of participants who initiate threads and those who responds to them.

Researchers should take advantage of the artifacts produced through NIC processes to better understand the implementation of NICs. Participants in NICs engage in many activities that result in written artifacts, such as root cause analyses, the development of a theory of action, a plan for implementing an intervention, summaries of progress-monitoring data, and summaries of conclusions and next steps. All three implementation studies reviewed analyzed NIC artifacts to better understand implementation. Researchers should consider collecting artifacts and developing rubrics to rate the artifacts for evidence of intended implementation and quality. The artifact review should indicate school completion of PDSA milestones (Implementation Question 3), alignment of these milestones with the NIC-level theory of practice improvement (Implementation Question 4), and extent to which schools are collaborating on specific PDSA milestones (also Implementation Question 4).

There is a need for new and better thinking about the best approaches to identify and study the outcomes of NICs. The rising popularity of NICs as an approach to improvement in education underscores the need for evidence that NICs are effective. To date, however, little research has been conducted on their outcomes. The four impact studies reviewed in this report focused on student outcomes, but it is not clear that the analyses answers the question “Did participation in the NIC result in improved student outcomes?” For three of the four studies reviewed, the NIC was only one component that factored into the development of the intervention.

To determine the causal effect of NICs on student outcomes, researchers should think carefully about what constitutes the treatment and comparison groups. This is not an easy task because, by design, NICs tend to have changing membership as new schools (and their teachers or administrators) join the NIC and other schools (and their teachers and administrators) leave it. NICs identify a problem of practice and develop and test an intervention meant to address that problem, but the intervention evolves over time and NIC membership may change, as well, as NIC members engage in the iterative process of testing and refining the intervention. Some of the questions that researchers need to grapple with include the following: “Is there is a certain point at which an intervention could be considered ‘final’ enough to test?” “Is assessing the impact of the ‘final’ intervention on student outcomes equivalent to assessing the impact of the NIC on student outcomes?” “How would we construct a comparison group to determine the causal impact of NICs on student outcomes?” Researchers will need to pay close attention to methodological issues, including power to detect a treatment effect and the threat of attrition. Finally, researchers should consider factors that may mediate the effect of participation in a NIC on student outcomes. NICs intend to change the way teachers do their work, which would be important to capture in a mediation model but would pose challenges related to measurement.

Researchers also should consider what other outcomes to measure in addition to student achievement. For instance, often NIC facilitators aim to build practitioners’ capacity to engage in systematic improvement on their own. Researchers could administer surveys to participants to assess their perceptions of their ability to facilitate NICs, including organizing and engaging in PDSA cycles, institutionalizing improvement science principles to address additional problems over time, and scaling up effective practices to improve student, school, or system performance. In addition, researchers should consider conducting retrospective studies to assess the sustainability of NICs over time, answering questions about the length of time practitioners were able to engage in improvement science methods with fidelity.

Limitations

Because NICs are an emerging approach to improvement in education, it is possible that the systematic search conducted using education research databases did not uncover all available evidence on the implementation and outcomes of NICs. There may be research on NICs published informally, such as on websites or through white papers that are not indexed by education databases. The review team did not conduct an extensive search of the gray literature to determine whether additional studies needed to be reviewed. In addition, as the field grows, we expect that additional research on the implementation and outcomes of NICs will emerge and that the search of education research databases will need to be updated at least every 6 months to uncover new research. Finally, there are many different approaches to improvement that have the same characteristics as NICs but are not explicitly referred to as NICs. Focusing on the research on NICs may lead to underestimating the research on NIC-like improvement processes.

Appendix A. Coding Template

Study-Level Information

Full citation:

Network Information

Improvement focus (e.g., math instruction, community college completion):

Improvement approach (e.g., Plan-Do-Study-Act cycles, inquiry cycles):

Level of network (i.e., school, district, state):

Number of participants in network:

Network setting (i.e., geographic location, urbanicity):

Implementation

If the article assesses implementation, code the following:

Research questions:

Sample information:

Research methods:

- Data and measures:
- Research design:

Findings:

Impact

If the article assesses impact, code the following:

Research questions:

Sample information:

Research methods:

- Data and measures:
- Research design:

Findings:

References

Ansell, C., Reckhow, S., & Kelly, A. (2009). How to reform a reform coalition: Outreach, agenda expansion, and brokerage in urban school reform. *Policy Studies Journal, 37*(4).

Bell, M., Jopling, M., Cordingley, P., Firth, A., King, E., & Mitchell, H. (2006). *What is the impact on pupils of networks that include at least three schools? What additional benefits are there for practitioners, organizations and the communities they serve?* Retrieved from http://www.curee.co.uk/files/shared/curee-nlc-systematic-review-impactofnetworkd.pdf

Bryk, A. S. (2015). (2014) AERA distinguished lecture: Accelerating how we learn to improve. *Educational Researcher*, *44*(9), 467–477. Retrieved from https://eric.ed.gov/?id=EJ1084855

Bryk, A. S., Gomez, L. M., LeMahieu, P. G., & Grunow, A. (2015). *Learning to improve: How America's schools can get better at getting better*. Cambridge, MA: Harvard Education Press. Retrieved from https://eric.ed.gov/?id=ED568744

Carnegie Foundation for the Advancement of Teaching. (n.d.) Building a teaching effectiveness network. Retrieved from https://www.carnegiefoundation.org/our-work/previous-improvement-work/bten/

Ell, F., & Meissel, K. (2011). Working collaboratively to improve learning and teaching of mathematics in a rural New Zealand community. *Math Education Research Journal, 23,* 169–187.

Evans, S. D., Rosen, A. D., Kesten, S. M., & Moore, W. (2014). Miami thrives: Weaving a poverty reduction coalition. *American Journal of Community Psychology, 53*(3–4).

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.

Huang, M., & Yamada, H. (2017). Maintaining success rates: Does Statway sustain its impact as it scales to new classrooms and institutions? Stanford, CA: Carnegie Foundation for the Advancement of Teaching.

Klein, A. (2018, May 8). "Continuous improvement" model woven into state ESSA plans. *Education Week.* Retrieved July 17, 2018, from https://www.edweek.org/ew/articles/2018/05/07/continuous-improvement-model-woven-into-state-essa.html

Rohanna, K. (2017). Breaking the "adopt, attack, abandon" cycle: A case for improvement science in K–12 education. In C. A. Christie, M. Inkelas, & S. Lemire (Eds.), *Improvement Science in Evaluation: Methods and Uses. New Directions for Evaluation, 153*, 65–77.

Russell, J. L., Bryk, A. S., Dolle, J. R., Gomez, L. M., LeMahieu, P. G., & Grunow, A. (2017). A framework for the initiation of networked improvement communities. *Teachers College Record, 119*(5), 1–36.

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*, 465–494.

Yamada, H., Bohannon, A. X., Grunow, A., & Thorn, C. A. (2018). Assessing the first two years' effectiveness of Quantway®: A multilevel model with propensity score matching. *Community College Review, 46*(3), 257–287.

Yamada, H., & Bryk, A. S. (2016). Assessing the first two years' effectiveness of Statway®: A multilevel model with propensity score matching. *Community College Review, 44*(3), 179–204.

Established in 1946, with headquarters in Washington, D.C., the American Institutes for Research® (AIR®) is a nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance, both domestically and internationally, in the areas of education, health, and the workforce. For more information, visit www.air.org.

MAKING RESEARCH **RELEVANT**

AMERICAN INSTITUTES FOR RESEARCH
1000 Thomas Jefferson Street NW
Washington, DC 20007-3835 | 202.403.5000
www.air.org

LOCATIONS

Domestic: Washington, DC (HQ) | Monterey, Sacramento, and San Mateo, CA | Atlanta, GA | Honolulu, HI | Chicago and Naperville, IL Indianapolis, IN | Waltham, MA | Frederick and Rockville, MD | Chapel Hill, NC | New York, NY | Columbus, OH | Cayce, SC | Austin, TX Arlington and Reston, VA | Seattle, WA

International: Algeria | Ethiopia | Germany | Haiti | Zambia

10462_01/20