Peabody Journal of Education

ISSN: 0161-956X (Print) 1532-7930 (Online) Journal homepage: www.tandfonline.com/journals/hpje20

Bringing Scientific-Professional Learning Communities into Practice: Exploring Variation in Educational Improvement Network Health and Development

Jennifer Lin Russell, Anthony S. Bryk & Jennifer Zoltners Sherer

To cite this article: Jennifer Lin Russell, Anthony S. Bryk & Jennifer Zoltners Sherer (2025) Bringing Scientific-Professional Learning Communities into Practice: Exploring Variation in Educational Improvement Network Health and Development, Peabody Journal of Education, 100:1, 64-81, DOI: 10.1080/0161956X.2025.2444843

To link to this article: https://doi.org/10.1080/0161956X.2025.2444843

Published online: 12 Feb 2025.

PEABODY JOURNAL OF EDUCATION
2025, VOL. 100, NO. 1, 64–81
https://doi.org/10.1080/0161956X.2025.2444843

Routledge
Taylor & Francis Group

Bringing Scientific-Professional Learning Communities into Practice: Exploring Variation in Educational Improvement Network Health and Development

Jennifer Lin Russell[a], Anthony S. Bryk[b], and Jennifer Zoltners Sherer[c]

[a]Vanderbilt Peabody College, Nashville, TN; [b]Carnegie Foundation for the Advancement of Teaching, Palo Alto, CA; [c]University of Pittsburgh, Pittsburgh, PA

ABSTRACT

The Improvement Network Health and Development Framework (INHD Framework) sets a vision for a complex new organizational form for practical problem solving: the Networked Improvement Community (NIC). One critical test of this idealized framework is to explore its usefulness in describing variation in these deliberatively formed, temporary problem-solving networks as they are operationalized in the public school context. This article explores the health and development of 34 improvement networks catalyzed by the Gates Foundation's Networks for School Improvement (NSI) initiative using evidence generated from a survey-based measurement system. We present evidence that at least six NSIs from the sample of 34 have normative practices and member attitudes that strongly align with the INHD Framework, suggesting that some intentional improvement communities aligned with the framework have come into existence, while another subset of five NSIs were significantly struggling to realize the idealized vision of a scientific-professional learning community. Following these high-level patterns, we examine evidence from program documentation to cross validate patterns found in network health and development identified through the survey.

Interorganizational improvement networks have proliferated in the education field as educators, system leaders, and various reform actors have become increasingly dissatisfied with the results produced by reforms dependent on local implementation of externally-generated programs and policies. Rising demands on local educational systems to both expand access to and improve the quality of educational offerings in the past several decades have strained the capacity of local educators and system leaders (Peurach et al., 2019). There is a critical need to build local capacity so that educators in schools and systems can engage in the problem solving, systemic improvement, and innovation necessary to meet our current societal aspirations for mass schooling. As a result, education leaders and reformers are looking to novel ways that expand local capacity and expertise.

The proliferation of improvement networks in the educational field received a major boost when the Bill & Melinda Gates Foundation launched its Networks for School Improvement initiative (NSI) in 2018. The NSI initiative provided competitive funding for hub organizations that wanted to operate a network of 10 or more middle and/or high schools working together to "advance high school graduation and college success for their Black and Latino students and students experiencing poverty" (Bill & Melinda Gates Foundation, n.d.). The initiative supported the launch of over 40 networks

CONTACT Jennifer Lin Russell jennifer.russell@vanderbilt.edu Vanderbilt Peabody College, 230 Appleton Place, PMB 414, Nashville, TN 37203-5721

across 23 states, which engaged over 600 schools serving more than 150,000 students in focused improvement efforts.

The NSI initiative represents a rich context to explore improvement network development. The Improvement Network Health and Development Study (INHD Study) works to conceptualize and measure the health and development of improvement networks. The project is grounded in the Improvement Network Health and Development Framework (INHD Framework) articulated by Russell et al. (2025, this issue) that outlines a vision for a complex new organizational form for practical problem solving: the Networked Improvement Community (NIC) operating as a scientific-professional learning community (Bryk et al., 2011, 2015). One critical test of this idealized framework is to explore its usefulness in describing variation in these deliberatively formed, temporary problem-solving networks as they are operationalized in the public school context. Drawing on the survey-based measurement system described by Bryk et al. (2025, this issue), this paper explores variation in the health and development of 34 networks catalyzed by the initiative.

We present evidence that at least six NSIs developed normative practices and member attitudes that strongly align with the INHD Framework, suggesting that intentional improvement communities consistent with the framework have come into existence in this initiative. These six networks illustrate that a robust improvement network can form around different improvement goals and be led by different types of hub organizations. As expected, the majority of NSIs ($N = 23$) had mixed developmental profiles with strengths in some dimensions of network health but not others. We also identified a subset of five NSIs that have clearly struggled to develop along lines consistent with the INHD Framework. Following these high-level patterns, we examine evidence from program documentation to cross-validate the patterns in network health and development identified through the survey results.

Conceptual framework

Improvement network health and development

Our study is grounded in the Networked Improvement Community concept (NIC) that describes an organizational form aimed at building a technical and social infrastructure to support collaborative, continuous improvement. NICs create opportunities for educators to engage in rigorous testing of practice changes, work collaboratively with colleagues within and across organizations, and accumulate practical insights that can yield substantive improvement in the educational processes that shape student learning and development (Bryk et al., 2011, 2015; Barron et al., 2024; Hannan et al., 2015; Russell et al., 2017, 2025, this issue; Yamada et al., 2018).

We conceptualize NICs as comprised of interdependent social and technical processes that enable operation as a scientific-professional learning community. They function to: (a) facilitate collaborative learning; (b) activate professionals to build practical knowledge; and (c) utilize disciplined, collaborative, and inclusive approaches to knowledge production and use (see elaborated discussion of scientific-professional learning communities in Russell et al. (2025, this issue). Drawing on theory and prior research, we developed the INHD Framework to describe a high-functioning NIC, what we refer to as a healthy improvement network. The framework posits six domains, when functioning effectively together, that constitute healthy improvement networks: hub leadership, network roles and engagement, continuous improvement processes, within-team connections, cross-team connections, and network culture (see Figure 1). The research and theoretical grounding for the framework are explicated in the first article in this themed issue (Russell et al., 2025, this issue). Here, we briefly outline these domains to ground the empirical findings we present in this article.

NICs are led by a central "hub" with responsibility for their development (Peurach et al., 2025, this issue). Hub leadership is viewed as the driver in the INHD Framework because the strategic actions of hub leaders catalyze the formation and operation of improvement networks (Lochmiller & Karnopp, 2024; Peurach, 2016; Peurach, 2025, this issue).

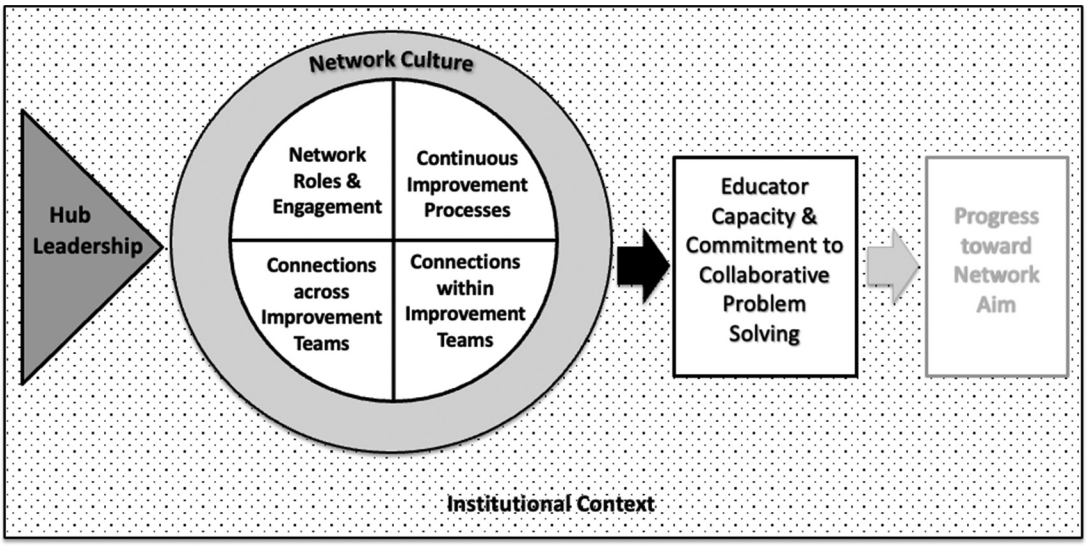

Figure 1. Networked improvement community health and development framework.

Four dimensions of the INHD Framework describe key components of the collaborative improvement process NICs aim to establish: network roles and engagement, continuous improvement, connections within improvement teams, and connections across improvement teams. NICs scaffold the engagement of members by defining roles that structure educators' participation in continuous improvement processes (Biag & Sherer, 2021). Continuous improvement processes include a set of practices that enable improvement teams to understand the problem they are trying to solve, engage in disciplined inquiry, and ground their work in relevant data and evidence (Bennett & Provost, 2015; Langley et al., 2009). The social structure of a NIC includes the formation of intentional connections within and across organizations that facilitate collaborative inquiry and the spread of promising practices across contexts (Burt, 2004; Goldsmith & Eggers, 2005; Weber & Khademian, 2008). The improvement process is supported by a network culture that promotes a collective identity in pursuit of an improvement aim and the use of evidence to ground improvement processes (Russell et al., 2025, this issue). In the short term, healthy networks build educator capacity and commitment to collaborative problem solving (Perlman et al., 2025, this issue), and over time produce progress toward the network's improvement aim.

The INHD Framework also recognizes that improvement networks operate in complex institutional contexts that shape network operations and member experiences (Yurkofsky et al., 2020). In healthy networks, district and school leaders prioritize, support, and value the work of the network (Poortman et al., 2022). When institutional contexts and network activities are not aligned, members experience challenges such as finding time to participate and integrating improvement activities into their professional responsibilities.

Developing healthy networks in complex environments

The INHD Framework outlines an ambitious set of technical functions and social structures. Actualizing this novel organizational form in interaction with public school systems is likely to confront implementation barriers. As hub leaders establish these temporary organizations, they draw members from existing organizations and seek to influence work in existing systems that operate in turbulent environments with plural goals. This often places great stress on educators' capacity, commitment, and time (Cohen et al., 2018; Larbi-Cherif et al., 2024; Peurach

et al., 2019, 2025, this issue; Woulfin & Spitzer, 2023). Hub leaders operating with limited authority over network members are challenged to structure participation in ways that are feasible given educators' limited time, that builds educator capacity to engage in improvement work, and that recognizes that this work is outside the scope of their prior training and largely counter-normative. As a result, we would expect a high degree of variability in the extent to which networks are able to realize the vision of a robust scientific-professional learning community. Given these formidable implementation challenges, an essential question is whether these social and technical problem-solving organizations can come to life in public school settings.

Conditions that may influence the formation of healthy networks

Given expected variation in the health and development of networks due to implementation challenges, another important question is whether certain design features and structural characteristics of networks may be associated with their capacity to launch and operate as healthy scientific-professional learning communities.

Given the hypothesized centrality of hub leadership, we would expect that hub capacity would be an important factor in the development of a NIC (Peurach et al., 2025, this issue). Hub capacity is in part a function of the kinds of expertise that hub organizations bring to the work. Hubs staffed by intermediary organizations, operating outside of traditional educational systems, may be able to mobilize novel expertise necessary to operate an improvement network (Cooper & Shewchuk, 2015). School districts may be able to staff hubs with personnel who are knowledgeable about the resources and needs embedded in local communities and have systemic authority to shape organizational goals and improvement activities. Universities may have specialized content expertise and greater analytic capacity to support data-based improvement work and knowledge of evidence-based practice. Hub efficacy is also a function of the specific practices and work routines utilized to lead networks, which in turn depends on the knowledge, skills, and professional capacities of its staff and prior experience in leading such efforts (Greany & Kamp, 2022; Peurach, 2016; Peurach et al., 2025, this issue).

Another factor that may shape the formation of healthy networks is the complexity of the outcomes they aim to produce. Domains of network health such as improvement processes and team connections develop in interaction with the goals pursued. For example, networks aiming to improve students' literacy or mathematics learning must address the complexity of the instructional process, one that is notoriously difficult to influence (Cobb et al., 2020). Keeping students on track for graduation is facilitated by the introduction of novel routines that have schools tracking real time data that predicts when students are getting "off track" and intervening with customized supports (Allensworth & Easton, 2007). Processes such as students' transition to college and career require collaboration among school and non-school actors such as counselors, college admission officers, and families. Thus, we would predict that networks with different improvement aims confront variable challenges in developing the necessary social and technical infrastructure to make progress on their aim.

As temporary organizations seeking to catalyze novel work processes, improvement networks likely benefit from stability in hub and network membership. Hub leaders need to induct members into new roles and practices characteristic of a scientific-professional learning community (Russell et al., 2017, 2025, this issue). This process includes developing educators' mindsets and capacities to engage in disciplined inquiry and the formation of connections within and across school teams (Biag & Sherer, 2021; Hannan et al., 2015; Joshi et al., 2021). These components of network development are more challenging when there is turnover in network members (Coburn et al., 2012; Holme & Rangel, 2012; Sutcher et al., 2016).

Methodology

In this article we begin to explore network health and development in the NSI initiative by analyzing the results from administration of the Improvement Network Health and Development Survey (INHD Survey) in spring 2023. We explore three primary questions:

- Do NSIs vary in their level of development as indicated by the 2023 INHD Survey administration results?
- What factors are associated with variation in development?
- Do survey-based patterns in developmental variation align with hub leader accounts in program documentation?

We first utilized network member responses to the 2023 INHD Survey to characterize the range in network development that had emerged by the end of the fifth year of the NSI initiative. We then explored potential explanations for developmental variation, including the alignment with local contexts and the structural characteristics of the networks. Finally, we explored whether the self-reported practices utilized by hub leaders corroborate characterizations of some networks as more developed and others as less developed. In the following sections we outline key methodological choices that enabled this exploratory investigation.

INHD Survey (measures & administration)

Grounded conceptually in the INHD Framework, our project team developed a set of measures from the INHD Survey to assess network member perceptions of the functioning of a scientific-professional learning community. The measures were drafted or adapted from prior surveys in dialogue with the framework. More details about survey development and psychometric properties are available in the second article of this themed issue (Bryk et al., 2025). In brief, there are 19 core survey measures, each aligned to specific domains in the INHD Framework. For example, the internal team connections domain has three core measures: team norms, processes and support, and collaborative inquiry.

We administered the INHD Survey to all members of each NSI annually. Typically, this occurred in each network during their last network-wide convening of the year (between March and May). Surveys were administered through an online portal. The average response rate across networks for the 2023 administration was 63.6%, with individual network response rates ranging from 47.5% to 84.6%.

Exploring network variation

As described in the second article in this themed issue (Bryk et al., 2025), a Rasch Rating Scale was created for each measure based on the responses to the survey in spring 2021, the first year in which all networks participated. We computed an empirical Bayes estimate of the average response on each measure for each network that year. Based on the distribution of these network estimates in 2021, we created cutpoints for each measure at the 25^{th} and 75^{th} percentiles. These cutpoints served as a way to differentiate network development. As the cutpoints established in 2021 remain constant for every study year, it is possible for more or fewer networks to wind up in the top or bottom quartile in any year. The results presented in this article apply these cutpoints to the spring 2023 survey responses.

Using the 2023 empirical Bayes estimates, we scored each measure of network health for each network as + 1 (top quartile using 2021 cutpoints), 0 (middle 50% using 2021 cutpoints), and −1 (bottom quartile using 2021 cutpoints). We then averaged these scores across the set of measures within a domain to create a *domain* score for each network on each of the six core domains (hub leadership, continuous improvement, network roles and engagement, internal team connections, cross team connections, and

network culture). Lastly, we averaged these six domain scores to create an *omnibus measure of network health* for each network.

The variation in these omnibus network health scores allowed us to identify a clear subset of *well-developed networks*. Six stood out as exemplary. Their omnibus scores were 0.68 or higher. This is out of a total possible score of 1.0 if a network were to have all measures in the top quartile. Among these six networks, there was only one negative domain score out of 36 possibilities. In contrast, five networks appeared to be struggling. These *least-developed networks* had omnibus scores that were −0.78 or lower. (If all measures were in the bottom quartile, a network's omnibus score would be −1.0.) Every domain score for every network in this least-developed group was negative.

Visualizing the variation between these two clusters

To illustrate the substantive difference between the well-developed and least-developed subsets of networks, we rely on the probability density displays which are designed to display the variation in responses among network participants on a given measure. (For more on these visualizations, see the second article in this themed issue: Bryk et al., 2025.) In a network with strong normation, we would expect to see 80% or more responses in the two most positive response categories (e.g., suggesting network members generally strongly agreed and agreed with the items constituting a measure) and less than 20% in the two least positive (e.g., slightly agree and disagree). In contrast, in a newly formed network, or one struggling to actualize as an intentional community, responses would typically be more varied. By varied we mean that although some participants may offer very positive responses, many others offer neutral to negative responses.

Analysis of variation in structural characteristics of networks

We then explored how networks at the top and bottom end of the distribution (well-developed versus least-developed) varied with respect to a series of structural characteristics that might be expected to influence network development based on our literature review. The data on structural characteristics came from documentation collected by the Gates Foundation, including funding proposals and annual progress reports submitted by hub leaders. We generated information about network membership from network rosters submitted by hub organizations each year as part of the survey administration process; the rosters provide data on network size, average tenure (years in the network) of school-based and hub members, and member and hub stability. Given the small number of networks in each cluster, we rely on a descriptive presentation of results.

Analysis of archival documents to triangulate survey-based differences between clusters

We employed qualitative analyses of program documentation to explore more nuanced aspects of network development. These analyses aimed to triangulate evidence to validate the differences in network health and development between the well-developed and least-developed clusters of networks (identified through analysis of network members' survey accounts) and provide additional insights into the specific practices employed in the networks and the implementation challenges they encountered.

The foundation collected extensive documentation from each hub organization about their annual activities. At the end of each funding year, hub leaders wrote a report following a standard template that prompted reflection on whether networks had achieved annual goals, factors contributing to their annual progress (or lack thereof), and goal setting for the upcoming year. Hub leaders provided considerable detail about annual network activities; for example, the documents included detailed descriptions of the agendas for network convenings and the way inquiry cycles were organized and supported. Reports provided evidence to support claims that the network was (or was not) making expected progress toward improvement goals, with attention to potential explanations for the level of

progress. Each annual report tended to be in the 20-to-30-page range. While these data sources have some limitations given their function as an annual report to a funder (which risks lack of transparency and a positive bias), the level of detail and the close ongoing interaction between program officers and networks made us confident that these were a valid source of data about network leaders' strategies and actions in operating the networks.

After first reading and annotating program documentation, we developed a matrix to structure within-case analyses (Miles et al., 2014) for the well-developed and least-developed NSIs that included program years as columns and the core dimensions of the INHD Framework as rows (e.g., continuous improvement processes, connections within improvement teams, etc.). Based on a second review of each annual report and initial jottings, we summarized network activities associated with each framework dimension for a given year. For example, a summary for 2023 in the "connections across improvement teams" dimension would outline the way the NSI structured opportunities for educators to meet and collaborate with network colleagues from other schools. Overall, we were able to populate rich and detailed matrices for each network for all domains except network culture. Given the nature of the annual reports, we could identify very limited information that commented on activities associated with building a supportive network culture, and therefore we removed culture from the matrices and note this as a limitation of our exploratory analyses.

After completing the within-case analyses for each network, we wrote memos for each network health domain that detailed the associated network activities undertaken by each of the eleven focal networks. We explored differences between the most- and least-developed networks and summarized these insights into a cross-case meta-matrix (Miles et al., 2014).

We also engaged in inductive coding to explore the barriers that NSIs were encountering as they launched and operated their networks. This process identified implementation challenges such as hub capacity, local initiatives that compete for educator attention, and school-based staffing shortages. After identifying a comprehensive list of barriers encountered by the eleven case networks, we developed a cross-case matrix that identified which networks reported each barrier. This allowed us to explore whether the most and least-developed networks faced different implementation challenges.

Analysis of open-ended survey responses to triangulate structure-agency finding

During our analysis of the progress reports written by hub leaders of the six well-developed networks, we noticed a theme: as they developed the work in their respective networks, hub leaders were balancing considerations about organizational structure and standard routines with creating space for individual member agency. This theme was present in passages describing the need to create specific structures that organized members' engagement in collaborative improvement work, while also providing opportunities for educators to exert individual creativity and responsiveness to local needs. As we began to attend to this theme, we went back and read the progress reports of networks in both clusters that we had a previously reviewed and gained confidence that it was more prominent in the rationales for network activities written by well-developed NSIs than the least-developed NSIs.

To further examine this emerging theme, we turned to member responses to open-ended survey questions that asked members about the benefits and challenges of network participation. We marked passages that addressed structure and/or agency themes and subsequently explored differences in the responses from members in well-developed and least-developed networks.

Findings

The INHD Survey characterizes distinct clusters of networks

Out of the sample of 34 networks, a subset of six networks displayed a distinct pattern of survey responses: their members consistently offered strong positive endorsements on the core network health measures including strength in hub leadership, network roles and engagement,

continuous improvement processes, network connections across improvement teams, and network culture. This pattern of member responses provides initial evidence that these six networks formed as cohesive and intentional communities. Based on members' survey responses, we refer to these six networks as "well-developed." In contrast, five networks out of 34 displayed consistently more variable and less positive member responses on the core network health measures; we refer to these five networks as "least-developed." The remaining 23 NSIs displayed a mixed profile with some dimensions of network health more developed and others less developed. The presence of six well-developed networks and five least-developed networks in the NSI portfolio provides evidence that the survey helps us distinguish networks at different stages of health and development. In this initial exploration, we focus on the six most-developed and five least-developed NSIs to better understand the conditions that support or inhibit network development as educators have tried to bring the scientific-professional learning community concept into practice. In this section of the article, we focus on the differences between the most and least developed networks; in later sections, we present evidence generated from our analysis of qualitative data to support this distinction.

Figure 2 illustrates the starkly different pattern in survey responses exhibited by members of the most-developed networks compared to the least-developed networks. As an example, we spotlight member responses to two hub leadership measures that are indicators of the role hub leaders play in catalyzing the social and technical work of a scientific-professional learning community: (1) *relational trust* with hub leaders; and (2) member views about hub leaders' *management of the knowledge resources* necessary to advance improvement work. See the first two probability density displays in Figure 2. In both areas, members in well-developed networks were more consistently positive in their perceptions of hub leadership as evidenced by the probability density displays. In the well-developed networks, nearly 90% of members strongly agreed or agreed with items indicating they have confidence in the expertise and management of networks leader (relational trust); in the least-developed networks, just over 60% of members strongly agreed or agreed. Given the central role hub leaders play in the social structure of a scientific-professional learning community, it is important that network members experience leaders as trustworthy and respond favorably to the leadership they provide in structuring member engagement in the network (Russell et al., 2025, this issue). Similarly, network members in well-developed networks were much more likely to strongly agree or agree with items indicating that hub leaders are consolidating the learning generated by improvement teams and supporting its spread and use in other contexts. Again, this pattern is consistent with our conceptualization of a scientific-professional learning community and the role hub leaders play in identifying promising practice changes based on evidence from inquiry routines and leveraging trusting relationships to entice other network members to try changes that have evidence of efficacy in other contexts (Russell et al., 2025, this issue). Both strong relational trust and quality knowledge management can be viewed as normatively held within the most developed networks as less than 15% of members held contrary views.

Challenges catalyzing connections within teams

There was only one domain of network health for which the most-developed six networks did not demonstrate substantially stronger member responses: the measures that examine members' interactions with colleagues within their improvement teams (typically school-based teams). Although slightly more positive, responses on these measures in the well-developed networks varied along lines similar to the least-developed networks. One notable difference between groups is that over 30% of members in the least-developed networks never engaged in collaborative inquiry with their teams, as opposed to just under 20% of members in the most-developed networks. Overall, this finding is unexpected and speaks to the general challenges to catalyzing strong connections within school-based improvement teams. It also suggests that even the most-developed networks have not yet fully realized this aspect of a scientific-professional learning community.

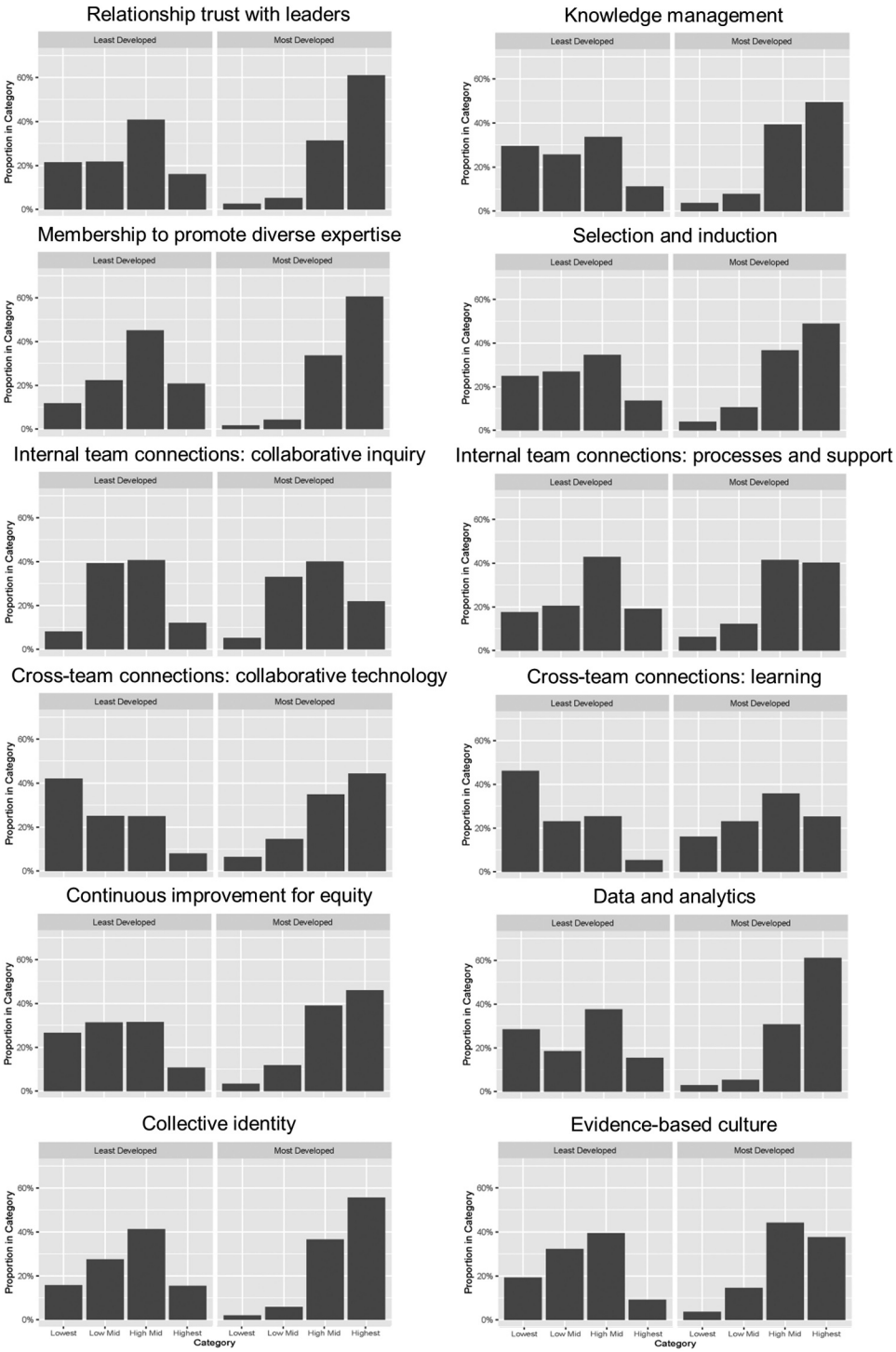

Figure 2. Comparison of network health and development measures between well-developed and least-developed networks.

Initial exploration of differences between well- and least-developed networks

Having used survey evidence to establish that one cluster of networks appears to be well-established and another appears to be less-developed, we briefly explore a range of potential explanatory factors including the role of local contexts for improvement and networks' structural and design features. Although our research design does not support causal inferences, this exploration lays the groundwork for future research.

The role of local contexts for improvement

Improvement networks are a new organizational form introduced into existing contexts of schools and districts. We hypothesize that the likelihood of this network structure taking deep root depends on these local context conditions (Russell et al., 2025, this issue). To explore this, we compared participant reports about their context conditions in the well-developed and least-developed networks. These results are presented in Figure 3.

Network members in both clusters report confronting a significant number of challenges with finding time for participation, receiving the support necessary to do the work, and integrating the improvement activity into their work. Although the incidence of these challenge reports is somewhat less in the well-developed networks (i.e., 30% versus 38% in the fewest challenges category), the differences here are not as large as seen in other network measures. Networked improvement is new

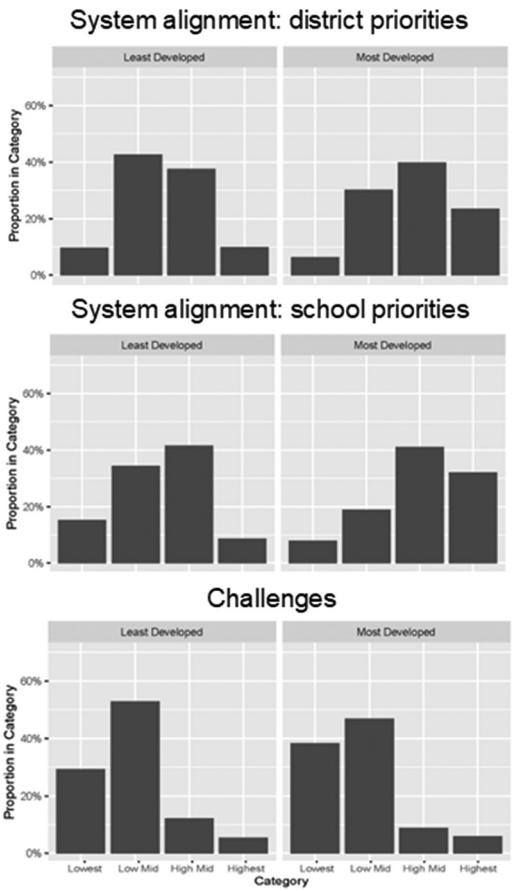

Figure 3. Probability density displays for network challenges and context measures.

work for all involved. Time is the scarcest resource in schools, and freeing teachers to take up improvement efforts is challenging (Woulfin & Spitzer, 2023).

In contrast, significant differences emerge between the two clusters on the questions that probed the alignment of the improvement efforts with district and school priorities. See Figure 3. The strongest of these differences exist at the school-site level. Only 26% of respondents in well-developed networks reported concerns here (i.e., responding in the two most negative response categories). In contrast, in the least-developed networks, over 50% viewed (mis)alignment between their school's priorities and their work in the network as a problem. These findings are consistent with an extensive body of research on how the organization and leadership at a school site affects the uptake and quality of implementation of reform initiatives (Bryk et al., 2010; Coburn, 2006). Although these contextual differences are not deterministic of success, this evidence suggests that aspects of context may well influence the ability of a network to take root in a given setting.

Network structural features

We found differences in several basic network characteristics between the two clusters. See Table 1. Four out of five networks in the least-developed cluster are focused on instructional improvement; the fifth network aimed to improve pathways to postsecondary education. Likewise, four of the five least-developed networks operate in single large urban districts. Although neither of these patterns is definitive, these results suggest that specific challenges may be associated with launching instructionally focused networks in these contexts.

On average, networks in the most-developed cluster were in operation longer: an average of 4.1 years in the most-developed cluster and 2.9 years in the least-developed cluster. This difference suggests that catalyzing a well-developed scientific-professional learning community takes time.

There was more hub turnover among the least-developed networks, with the well-developed groups either having stable hubs ($N = 3$) or some but not significant hub turnover ($N = 3$). In contrast, only one of the least-developed networks had stable hub leadership and two experienced high levels of hub turnover. In fact, one of the least-developed networks experienced an almost complete turnover in the hub team. Similarly, the well-developed networks had more stability in network members: the well-

Table 1. Structural characteristics of most-developed and least-developed networks.

	Imp Goal	Type of Hub Org	Launch Date (and cohort)	~Members in 2023	Districts	Member Stability (%>1 yr)	Avg. Tenure (Yrs) of 2023 Members	Hub Stability
MD1	Early Warning	NGO	F 2018 (1)	50	Multiple	88%	2.65	Stable
MD2	Post-Sec Match	NGO	F 2019 (1b)	300	Single	63%	1.89	Stable
MD3	Post-Sec Match	IHE	F 2018 (1)	200	Multiple	76%	2.33	Stable
MD4	Instructional	NGO	F 2020 (2)	60	Multiple	36%	1.66	Some Turnover
MD5	Instructional	District	F 2018 (1)	70	Single	68%	2.28	Early Turnover (then Stable)
MD6	Instructional	NGO	W 2021 (3)	65	Single	90%	2.68	Some Turnover
LD1	Post-Sec Match	IHE	W 2020 (2)	120	Multiple	57%	1.94	Some Turnover
LD2	Instructional	NGO	F 2020 (2)	40	Single	65%	2.13	Some Turnover
LD3	Instructional	NGO	W 2021 (3)	130	Single	80%	2.36	High Turnover
LD4	Instructional	District	F 2019 (1b)	120	Single	15%	1.2	High Turnover
LD5	Instructional	NGO	W 2022 (2)	60	Single	39%	1.39	Stable

developed networks averaged 70% of members in 2023 who had participated in the network for more than one year, while the least-developed networks averaged 51%. The average tenure of network members in the well-developed was 2.25 years compared to 1.80 years for the least-developed networks.

No differences between the two clusters, however, were found in terms of the type of hub organization. All three types of organizations leading NSIs, including system (districts) and nonsystem actors (e.g., IHEs, NGOs) were found in both clusters.

To more deeply explore differences between the two clusters, we turned to qualitative analysis of the evidence provided by these networks in their annual reports to the foundation. This analysis helped us corroborate and illustrate the survey-based differences we identified between the most-and least-developed networks.

Qualitative evidence lends support to survey-based identification of distinct clusters

As described in our conceptual framework, scientific-professional learning communities include both technical and social dimensions. Attention to structuring and supporting educator engagement in disciplined inquiry was common to both the most- and least-developed NSI clusters. Both invested in the development and use of tools and routines to support member engagement in continuous improvement processes related to their network aims, often with explicit attention to equity. These routines included structured inquiry cycles (e.g., Plan-Do-Study-Act cycles) and tools to scaffold engagement in cycles (e.g., workbooks, data collection templates, etc.), as well as attention to team processes that supported collaborative continuous improvement (e.g., protocols to guide conversations about data from improvement cycles). However, analyses of annual progress reports generated evidence that the well-developed networks in our sample exemplified greater attention to the integration of both technical and social processes around a shared working theory of improvement, suggesting that members in these NSIs may have experienced deeper supports for collaborative inquiry processes.

More sophisticated technical supports for improvement

The technical supports that hub leaders built in the most-developed networks were different in three specific ways from those in the least-developed networks. First, the most-developed networks maintained a very specific and iteratively refined working theory to guide their collective efforts. One of the most-developed NSIs created a visual representation of the "On Track (to Graduation) Roadmap," which anchors all work in the network. It illustrates how findings from root cause analysis, the network's three change packages, and corresponding interventions all contribute to increases in the proportion of minoritized students passing courses required for graduation. Three of the least-developed networks struggled to focus their work on specific processes or drivers associated with a targeted improvement aim. One of the least-developed NSIs shifted the focus of their improvement work entirely in its third year.

Second, the most-developed networks invested in the development and use of measures and/or data tracking platforms aligned to their working theory that enabled improvement teams to efficiently access evidence needed to engage in rigorous testing of practice changes. One network developed a college access dashboard that collects data from school portals (e.g., student FAFSA completion and other steps in the college access process) and a senior exit survey (measuring student perceptions of progress towards a college ready culture, school-level supports, and belonging). The network represents these data in ways that allow teams to dynamically track progress toward their annual aim. In contrast, the least-developed networks struggled to get a data infrastructure in place that supported team improvement work. For example, one instructional network took three years to arrive at a common focus for their efforts, which hampered their ability to identify measures that would be meaningful for all teams to track. Another network struggled to get access to timely information on students' post-secondary transition behaviors such

as enrollment in postsecondary options despite a multiyear effort to form partnerships with local colleges.

Finally, the most-developed networks introduced systematic methods to consolidate learning from testing cycles. For example, one NSI systematically visualized school-level trends on their leading outcome measures, allowing the hub to identify schools that were achieving significantly more positive results. They then reviewed documentation of these outlier schools' improvement work and interviewed team members to learn in-depth what they were doing and to identify evidence-supported practices that could be spread to other schools. Another well-developed network built a routine called "cycle reviews" to identify learning at the school level and leverage that learning to spread practice changes throughout the network. In four of the six most-developed networks, hub leaders built "change packages," which included an articulated theory of change, aligned change ideas, and measures for assessing the impact of change ideas. This depth of attention to these key technical processes was not evident in the least-developed networks' progress reports.

Greater capacities for social learning

The most-developed networks also displayed evidence of more attention to the social processes that support networked improvement than those in the least-developed cluster. One of the most-developed networks created affinity groups focused on specific components of their literacy achievement problem: improvement leads from multiple schools met weekly with their affinity groups to reflect on data collected through testing of practice changes in their respective schools. Other well-developed networks invested in the ongoing refinement of the collaborative routines within school-based teams. For example, in one NSI, improvement coaches from the hub conducted formative assessments of improvement team inquiry cycles and intervened to increase the quality of inquiry processes.

Progress reports indicate many of the NSIs in both clusters faced challenges convening network members, particularly in the COVID-19 context. However, some NSIs had better strategies to mitigate these challenges. For example, one of the least-developed networks reported they were finding it challenging to get improvement teams to meet due to staffing shortages that forced teachers to substitute for other teachers during their planning periods (when teachers would typically engage in collaborative improvement work). Some NSIs found ways to continue to meet despite these challenges. For example, one well-developed NSI working with schools in multiple districts adapted their strategy from holding a full network-wide convening twice a year to multiple convenings with subsets of districts. Another well-developed network hub redesigned their planned engagement with schools, realigning time allocated to site visits and face-to-face convenings toward a design for "interaction and tight facilitation in the on-line context."

The most-developed networks had intentionally designed coaching systems: hub staff designated as "improvement coaches" provided job-embedded supports to educators. For example, coaches supported schools in scaffolded improvement projects. Although the hubs of many of the least-developed networks had designated coaches, each struggled due to issues such as coach turnover, capacity, or late implementation of coaching. Several of the least-developed networks struggled to recruit sufficient improvement coaches to support schools. Most significantly, one of the least-developed networks lost all of their core staff except two coaches in the third year of operation; as a result, a hub member wrote that the fourth year of operation felt like, "a second Year 1 of the project—a restart."

Balancing standardized processes and individual agency

A final way the two clusters varied was in the ways they sought to balance standardized processes and individual agency. Members of a scientific-professional learning community are working together to solve a shared problem. Working in such a community requires some standardization of work; these shared routines, tools, and/or structures both support and constrain individual actions. For example, one well-developed network iteratively refined their "framework for collaborative, continuous improvement" and related tools used by coaches. They specifically noted that using the model aimed to reduce variability among coaches with respect to the support they provided to their teams.

Yet, at the same time, school-based teams were able to exert agency in the continuous improvement process by setting their own annual improvement priorities and selecting practice changes to test that aligned to these aims.

One of the well-developed networks referred to moving toward "tighter team cycles of inquiry" to advance quality disciplined inquiries, while at the same time allowing teams to select from an improvement menu that included possible practice changes. In another well-developed network, members' open-ended survey responses highlighted both the value of the structure the network provided and the individual agency afforded:

> *This careful measurement of intervention practice is rare outside the network, and it's exciting to me that because we've participated in it in such a structured, disciplined way, it is beginning to become a habit of mind.*
>
> *I value having a voice and being treated as a professional.*
>
> *I valued working with a group of people who are committed to elevating teacher voice and learning from the many different contexts in which teaching is done. I appreciate all of the effort that goes into learning how change can happen in real classrooms instead of it being mandated from above.*

Attention to both structured collective activity and individual agency was also represented in the way well-developed networks designed for social interaction. For example, hub leaders in one well-developed network created opportunity for choice within a defined structure during convenings:

> *One of the deliberate design decisions of the network this year was to incorporate more cross school collaboration . . . participants opted into one of the four change ideas (assessment and data review, equitable grading, student voice, EL supports), and they explored resources on the idea and shared approaches.*

In the least-developed networks, our findings suggest that network hubs may have either provided too much structure, contributing to limited participant buy-in or commitment, or too little structure, contributing an idiosyncratic character to the work. Network members' open-ended survey responses in the least-developed networks surfaced challenges associated with too little structure, such as:

> *Please train us in a few change ideas, then we can choose what we want to try. There is too much ambiguity in the process.*
>
> *Give ideas that work instead of us brainstorming all the time. Get ideas on how to involve parents, help parents complete their FAFSA, educate parents more.*
>
> *The meetings do not feel productive and are very indecisive. I feel more confused and don't feel it's a productive use of my time.*

In other least-developed networks, the responses suggested too much structure:

> *We felt stymied at times by the "official" (improvement) cycles.*
>
> *This work takes a considerable amount of time and isn't very flexible in terms of specific protocols. This makes it very challenging to implement the work in a way that feels meaningful.*
>
> *I am not convinced that the structure and actual logistics of this network facilitated this best practice in a meaningful way. Often it felt like having to complete the to-dos and deadlines actually interrupted our flow of instruction.*
>
> *I feel the overall premise of the work is slightly flawed in the sense that conversations continually moved toward a best practice that we could implement widely, while I feel that by definition the practice we were using needed to evolve continually based on varying circumstances.*

Discussion

The NSI initiative offered an exceptional opportunity to explore the development of improvement networks given the number of networks formed, the diversity of contexts involved, and the natural variation that emerged in the uptake of the Networked Improvement Community (NIC) concept. The INHD Framework (Russell et al., 2025, this issue), anchored in the notion of a scientific-professional learning community, enabled us to systematically examine the diversity in enactments displayed by these networks in practice. This descriptive information represents an important contribution to research on improvement networks, which to date has lacked detail as to how improvement networks operate (Poortman et al., 2022).

Our analysis of survey data identified six NSIs whose members offered consistently positive reports about the improvement practices and beliefs operating in their networks. This pattern of responses suggests that well-formed intentional communities have emerged in these networks. Identification of these NSIs as well developed was further corroborated by qualitative evidence suggesting these networks provided deeper supports for their members to engage in collaborative improvement work.

In contrast, there were five NSIs in the initiative that appear significantly less developed. Their members offered much more varied reports and, on average, weaker responses. Qualitative evidence indicates that these least-developed networks faced continuing challenges with the launch of their networks. Although well-developed networks confronted many of these same challenges, they were able to take more effective action to mitigate them. Notably, four of the least-developed NSIs aimed to catalyze instructional improvement and operated in large urban school districts. This finding is consistent with prior research that documents special challenges associated with launching large-scale change efforts on complex problems in the turbulent environments typically found in such district settings (Yurkofsky, 2022).

Taken together, these findings suggest the INHD Survey is useful for distinguishing between networks that are operating in more- or less-accord with the principles of a scientific-professional learning community.

Our exploration of factors associated with variation between the well-developed and least-developed networks is helpful in validating, corroborating, and exemplifying the survey results and lays the groundwork for future research. First, our findings suggest that catalyzing a vital scientific-professional learning community in education requires intentionally designed technical processes embedded in a social organization that coordinates the collective action of professionals. There is evidence that both clusters of networks were taking action to structure the technical processes and social connections theorized to support networked improvement. However, we found that the most-developed networks had achieved a deeper level of technical and social supports. Key components of this support include aligning continuous improvement work with an articulated theory of improvement and providing effective job embedded supports, such as coaching, for learning how to utilize continuous improvement approaches. These emerging findings begin to give a more nuanced descriptive portrait of the domains of effort we describe in the INHD Framework (Russell et al., 2025, this issue) and are consistent with prior research on the importance of coherence in school change efforts (Cobb et al., 2020; Fullan & Quinn, 2016; Honig & Hatch, 2004) and the power of coaching to support continuous improvement (Russell et al., 2020; Woulfin et al., 2023). The field would benefit from in-depth case studies that explore how hub leaders bring the core dimensions of the framework to life.

Second, our findings further suggest that when hub leaders are managing interdependent social and technical processes, challenges emerge in balancing the structure needed to socialize educators into new modes of collaborative problem solving and the agency necessary to foster educator motivation and commitment. Hub leaders in the NSI initiative varied in their capacity to manage this dialectic. The tension between structure and agency is consistent with prior research on network governance and operation. For example, Greany and Kamp (2022) note that networks enact governance and management structures to support network operations, but caution that these structures may

influence the extent that network members feel ownership for the work. Citing Hargreaves et al. (2015) and McCarthy et al. (2004), they observe that protocols and routines intended to support network impact need to be enacted in ways that prevent pushing the network toward "rule following" (p. 19). Our qualitative findings extend the INHD Framework by positing the need to jointly support the strategic management of structure and agency as a key dynamic in hub leadership, again a fruitful area for further investigation.

Third, an unexpected finding was that members in well-developed networks did not consistently experience more frequent and positive interactions within their school improvement teams than peers in the least-developed networks. This suggests that within-team connections may be harder for network hubs to influence because these teams are deeply embedded in school and district contexts and exist largely outside of their direct zone of control. This interpretation is consistent with research on professional learning networks more broadly that finds that school context factors, such as school leadership, shape the quality of team connections (Poortman et al., 2022). Further exploration is needed to better understand the challenges associated with catalyzing and sustaining vital school teams.

In this first look at a relatively large sample of improvement networks, we aimed to posit potential contributors and barriers to network development that can be explored further in subsequent research. The nature of our data sources affords both opportunities and limitations. The INHD Survey (Bryk et al., 2025, this issue) offers a first of its kind comprehensive assessment of the experiences of network members in a large sample of improvement networks. As an evidence source, the survey provides insight into the variable perceptions of network members regarding the functioning of their networks. Our qualitative exploration of network programmatic actions is based largely on documents written by hub leaders in their reports to a funder. Given the purpose of the documents, we worried at the outset that these accounts might offer a sanitized account of network activities. However, given that survey-based reports from members align substantively with these field-based records, this lends credence to a claim that meaningful differences operated in the work lives of members in these two different groups of networks.

In conclusion, this investigation adds new evidence to an emerging research base on the promise and challenges involved in forming educational improvement networks. Having established that the INHD Framework and Survey help us to see variability in the extent to which networks are operating in accordance with the principles of a scientific-professional learning community, an important next phase is to explore how patterns in network health and development are associated with valued outcomes, including members' perceptions of participation benefits and ultimately impacts on students.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This article was supported by funding from the Bill & Melinda Gates Foundation (INV-007724). Any opinions, findings, conclusions, or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the Foundation.

Notes on contributors

Jennifer Lin Russell is a professor of leadership, policy, and organizations at Vanderbilt University's Peabody College and codirector of the Improvement Scholars Network (with Donald Peurach). Her research focuses on organizing educational systems for improvement, with a particular emphasis on networked continuous improvement.

Anthony S. Bryk is president emeritus at the Carnegie Foundation for the Advancement of Teaching. He pioneered the Foundation's efforts to join the discipline of improvement science with the power of networks to create a new approach toward solving longstanding educational problems. His efforts are anchored in both academic scholarship on the sociology of organizations and practical experiences in improvement efforts in marginalized school communities anddistricts.

Jennifer Zoltners Sherer is a Research Associate at the University of Pittsburgh's Learning Research & Development Center and a member of the Partners for Network Improvement (PNI) team. She supports network leaders in their work to design, implement, and adapt improvement networks. She also serves as a hub leader for the STEM PUSH improvement network. Prior to receiving her PhD in Learning Sciences from Northwestern University, she was a teacher in Oregon.

ORCID

Jennifer Lin Russell http://orcid.org/0009-0002-7271-0709

References

Allensworth, E. M., & Easton, J. Q. (2007). *What matters for staying on-track and graduating in Chicago public high schools: A close look at course grades, failures, and attendance in the freshman year*. Consortium on Chicago School Research.

Barron, K. E., Hulleman, C. S., Hartka, T. A., & Inouye, R. B. (2024). Using a networked improvement community approach to design and scale up social psychological interventions in schools. *The Journal of Experimental Education*, *92*(2), 287–311. https://doi.org/10.1080/00220973.2023.2202832

Bennett, B., & Provost, L. (2015). What's your theory? *Quality Progress*, *48*(7), 36.

Biag, M., & Sherer, D. (2021). Getting better at getting better: Improvement dispositions in education. *Teachers College Record*, *123*(4), 1–42. https://doi.org/10.1177/016146812112300402

Bill & Melinda Gates Foundation. (n.d.). *Networks for school improvement*. Bill & Melinda Gates Foundation. https://usprogram.gatesfoundation.org/what-we-do/k-12-education/networks-for-school-improvement

Bryk, A., Li, A., Luppescu, S., & Bui, M. A. (2025). Examining the validity of practical measures of improvement network health and development. *Peabody Journal of Education*, *100*(1), 28–47. https://doi.org/10.1080/0161956X.2025.2444840

Bryk, A. S., Gomez, L. M., & 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.

Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. (2010). *Organizing schools for improvement: Lessons from Chicago*. The University of Chicago Press.

Burt, R. S. (2004). Structural holes and good ideas. *The American Journal of Sociology*, *110*(2), 349–399. https://doi.org/10.1086/421787

Cobb, P., Jackson, K., Henrick, E., & Smith, T. M. (2020). *Systems for instructional improvement: Creating coherence from the classroom to the district office*. Harvard Education Press.

Coburn, C. E. (2006). Framing the problem of reading instruction: Using frame analysis to uncover the microprocesses of policy implementation. *American Educational Research Journal*, *43*(3), 343–379. https://doi.org/10.3102/00028312043003343

Coburn, C. E., Russell, J. L., Kaufman, J. H., & Stein, M. K. (2012). Supporting sustainability: Teachers' advice networks and ambitious instructional reform. *American Journal of Education*, *119*(1), 137–182. https://doi.org/10.1086/667699

Cohen, D. K., Spillane, J. P., & Peurach, D. J. (2018). The dilemmas of educational reform. *Educational Researcher*, *47*(3), 204–212. https://doi.org/10.3102/0013189X17743488

Cooper, A., & Shewchuk, S. (2015). Knowledge brokers in education: How intermediary organizations are bridging the gap between research, policy and practice internationally. *Education Policy Analysis Archives*, *23*(118), n118. https://doi.org/10.14507/epaa.v23.2355

Fullan, M., & Quinn, J. (2016). Coherence making. *The School Administrator*, *73*(6), 30–34.

Goldsmith, S., & Eggers, W. D. (2005). *Governing by network: The new shape of the public sector*. Brookings Institution.

Greany, T., & Kamp, A. (2022). *Leading educational networks: Theory, policy and practice*. Bloomsbury Publishing.

Hannan, M., Russell, J. L., Takahashi, S., & Park, S. (2015). Using improvement science to better support beginning teachers: The case of the building a teaching effectiveness network. *Journal of Teacher Education*, *66*(5), 494–508. https://doi.org/10.1177/0022487115602126

Hargreaves, A., Parsley, D., & Cox, E. K. (2015). Designing rural school improvement networks: Aspirations and actualities. *Peabody Journal of Education*, *90*(2), 306–321. https://doi.org/10.1080/0161956X.2015.1022391

Holme, J. J., & Rangel, V. S. (2012). Putting school reform in its place: Social geography, organizational social capital, and school performance. *American Educational Research Journal*, *49*(2), 257–283. https://doi.org/10.3102/0002831211423316

Honig, M. I., & Hatch, T. C. (2004). Crafting coherence: How schools strategically manage multiple, external demands. *Educational Researcher*, *33*(8), 16–30. https://doi.org/10.3102/0013189X033008016

Joshi, E., Redding, C., & Cannata, M. (2021). In the NIC of time: How sustainable are networked improvement communities? *American Journal of Education*, *127*(3), 369–397. https://doi.org/10.1086/713826

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.

Larbi-Cherif, A., Glazer, J. L., & Ison, A. (2024). Continuous improvement in urban districts: Bringing environments back in. *American Journal of Education*, *130*(3), 363–394. https://doi.org/10.1086/729598

Lochmiller, C. R., & Karnopp, J. R. (2024). Understanding initiators' problem framing in the initiation of a networked improvement community. *Journal of Educational Change*, *25*(4), 1–27. https://doi.org/10.1007/s10833-023-09501-w

McCarthy, H., Miller, P., & Skidmore, P. (Eds.). (2004). *Network logic: Who governs in an interconnected world?* Demos.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE Publications.

Perlman, H., Bryk, A. S., & Russell, J. L. (2025). Measuring educators' perceived benefits of participation in improvement networks. *Peabody Journal of Education*, *100*(1), 82–99. https://doi.org/10.1080/0161956X.2025.2444844

Peurach, D. J. (2016). Innovating at the nexus of impact and improvement: Leading educational improvement networks. *Educational Researcher*, *45*(7), 421–429. https://doi.org/10.3102/0013189X16670898

Peurach, D. J., Cohen, D. K., Yurkofsky, M. M., & Spillane, J. P. (2019). From mass schooling to education systems: Changing patterns in the organization and management of instruction. *Review of Research in Education*, *43*(1), 32–67. https://doi.org/10.3102/0091732X18821131

Peurach, D. J., Jones, E. S., Duff, M., Sherer, J. Z., & Matthis, C. S. (2025). The practice and contexts of hub and district leadership: New directions in research on educational improvement networks. *Peabody Journal of Education*, *100*(1), 117–132. https://doi.org/10.1080/0161956X.2025.2444846

Poortman, C. L., Brown, C., & Schildkamp, K. (2022). Professional learning networks: A conceptual model and research opportunities. *Educational Research*, *64*(1), 95–112. https://doi.org/10.1080/00131881.2021.1985398

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. https://doi.org/10.1177/016146811711900501

Russell, J. L., Bryk, A. S., Peurach, D. J., Sherer, J. Z., Duff, M., Sherer, D., & Matthis, C. S. (2025). Catalyzing scientific-professional learning communities: A framework for conceptualizing the health and development of educational improvement networks. *Peabody Journal of Education*, *100*(1), 7–27. https://doi.org/10.1080/0161956X.2025.2444839

Russell, J. L., Correnti, R., Stein, M. K., Thomas, A., Bill, V., & Speranzo, L. (2020). Mathematics coaching for conceptual understanding: Promising evidence regarding the tennessee math coaching model. *Educational Evaluation and Policy Analysis*, *42*(3), 439–466. https://doi.org/10.3102/0162373720940699

Sutcher, L., Darling-Hammond, L., & Carver-Thomas, D. (2016). *A coming crisis in teaching? Teacher supply, demand, and shortages in the US*. Learning Policy Institute.

Weber, E. P., & Khademian, A. M. (2008). Wicked problems, knowledge challenges, and collaborative capacity builders in network settings. *Public Administration Review*, *68*(2), 334–349. https://doi.org/10.1111/j.1540-6210.2007.00866.x

Woulfin, S. L., & Spitzer, N. (2023). Time is ticking: The dynamics of education reform in the Covid-era. *Journal of Educational Administration*, *61*(3), 256–271. https://doi.org/10.1108/JEA-09-2022-0147

Woulfin, S., Stevenson, I., & Lord, K. (2023). *Making coaching matter: Leading continuous improvement in schools*. Teachers College Press.

Yamada, H., Bohannon, A. X., Grunow, A., & Thorn, C. A. (2018). Assessing the effectiveness of Quantway®: A multilevel model with propensity score matching. *Community College Review*, *46*(3), 257–287. https://doi.org/10.1177/0091552118771754

Yurkofsky, M. (2022). Environmental, technical, and representational uncertainty: A framework for making sense of the hidden complexity of educational change. *Educational Researcher*, *51*(6), 399–410. https://doi.org/10.3102/0013189X221078590

Yurkofsky, M. M., Peterson, A. J., Mehta, J. D., Horwitz-Willis, R., & Frumin, K. M. (2020). Research on continuous improvement: Exploring the complexities of managing educational change. *Review of Research in Education*, *44*(1), 403–433. https://doi.org/10.3102/0091732X20907363