Evaluation of the Networks for School Improvement Initiative—School-Level Implementation of Continuous Improvement

Appendices A–G

Michael S. Garet, Laura B. Stein, Ryan C. Eisner, Matthew J. Farmer, Sara Mitrano, Damon Blair, Emma Wilson, Katheen T. Jones, and Kianna Medina

FINAL REPORT, APRIL 2026

American Institutes for Research® | AIR.ORG

AIR®
Advancing Evidence.
Improving Lives.

Contents

Exhibits

Appendix A. Glossary

This glossary provides definitions of key terms pertaining to networks and continuous improvement (CI) as they are typically used in the Networks for School Improvement (NSI) initiative.

Aim statement. An aim statement articulates the goal of an improvement effort. An aim states what network participants are trying to accomplish. An aim statement should clearly specify how much, for whom, and by when.[1] Aim statements can be generated at various organizational levels (e.g., team-specific or networkwide).

Artifact. An artifact is a document that an intermediary, network, or CI team generates during CI work. For example, documents may include root cause diagrams or templates that teams used to plan inquiry cycles.

Change idea. A change idea is a specific practice or intervention that a CI team tests during inquiry cycles. Change ideas are typically designed to meet the goal outlined in the aim statement, focusing on the drivers that guide the network's theory of improvement.[2]

Change idea topics. The evaluation classified the topics that change ideas focused on using the following taxonomy:

- **Curricular changes.** Changes to the material taught in the classroom, including developing teacher content knowledge and tailoring curricula to student needs.
- **Shifts to pedagogic practices.** Changes to the way students and teachers engage with the academic material in the classroom, including individual and small-group supports provided within the classroom instruction context.
- **Changes in classroom culture.** Behavior management in the classroom context, establishing a positive culture and maintaining high expectations.
- **Social-emotional learning.** The development of personal and interpersonal skills beyond academic content, such as self-awareness, self-management, responsible decision making, relationship skills, and social awareness.
- **Academic advising and tutoring.** Outside the classroom instruction context, providing individualized or small-group support for students to improve their academic skills and grades, support timely progression through middle- and high school, and facilitate the transition to postsecondary education.

[1] Carnegie Foundation for the Advancement of Teaching. (2026, February 9). *Learning to improve glossary*. https://www.carnegiefoundation.org/resources/learning-to-improve-glossary

[2] For the purposes of CI artifact coding, a change idea is defined as a specific change in practice that can be tested over a short time frame.

- **Adult–student relationships and school culture.** Building relationships with students through mentoring and/or community-building efforts. The primary focus of relationship development is often on nonacademic content but could extend to academic material outside classroom instruction.
- **Family connections.** Changes that focus on family involvement unrelated to college application and enrollment, such as communicating with families about student progress or getting families involved in school governance.
- **Identifying students in need of academic support.** Using data to identify students who are struggling academically or otherwise need extra support; monitoring improving grade point average (GPA), course completion, and other academic indicators; building cross-functional teams that share data.
- **College access and affordability.** Assisting students with the college application and enrollment process, including Free Application for Federal Student Aid (FAFSA) completion.
- **College-going culture.** Promotion of a college-going mindset.
- **School systems and policies.** Redesigning policies, practices, and materials for equity.
- **Other.** Use only if the change idea does not relate to any of the topics above. Examples include building CI team capacity or opportunities for educator self-reflection and development.

Cohort. A cohort is a group of NSI grants that the Gates Foundation awarded around the same time. The foundation awarded the NSI grants in three cohorts: Cohort 1 grants were awarded in 2018, Cohorts 1B and 2 grants were awarded in 2019, and Cohort 3 grants were awarded in 2020.

Continuous improvement (CI). CI is a process in which practitioners engage in iterative cycles of inquiry by defining local problems of practice, testing potential change ideas, studying the results, and improving those change ideas.

Continuous improvement team (CI team). A group of educators that engages in CI (e.g., conducts root cause analysis and disciplined inquiry cycles) to address a local problem of practice.

Drivers. Drivers are factors that influence or produce more equitable student outcomes. The primary drivers are the major factors thought to lead to improvement. The secondary drivers are the interventions focused on the primary drivers.[3] Drivers are often shown as part of a diagram that illustrates the relationship between the change ideas, drivers, and aim.

[3] Bryk, A. S., Gomez, L., & Grunow, A. (2011). *Getting ideas into action: Building networked improvement communities in education.* Carnegie Foundation for the Advancement of Teaching. https://www.carnegiefoundation.org/resources/publications/getting-ideas-action-building-networked-improvement-communities-education/

Educational equity. Educational equity means providing students with resources, experiences, and environments—allocated based on circumstances and needs—so that students have equal access to opportunities for success. One of the major goals of the NSI initiative is to promote educational equity for Black students, Latino students, and students experiencing poverty. Intermediary organizations and the CI teams they support were charged with applying an equity lens to all CI processes, such as the setting of aims and the development of change ideas. The equity framework used for the NSI evaluation consists of four dimensions of equity (developed from Gutiérrez, 2009)[4]:

- **Access.** Providing access to teachers who are attentive to student needs (e.g., addressing teacher mindsets), access to supplies (e.g., equipment, software, materials in multiple languages) for effective learning, access to classroom environments that encourage student participation (e.g., perhaps using pedagogical tools that consider how students learn differently), and access to support or opportunity (e.g., tutoring, after-school programs for outside learning).
- **Achievement.** Enabling students to achieve traditional outcomes, including grades, test scores, course taking, graduation, and postsecondary attendance; and documenting outcomes and disaggregating data to identify gaps.
- **Identity.** Resources and interventions that are attentive to a student's or teacher's background, which can include personal characteristics; family and community histories; and their membership in social groups based on race, ethnicity, gender, class, status, ability, sexual orientation, religion, language, and many others. Identity situates education as a cultural practice in which learners can see themselves and others favorably.
- **Agency.** Resources and interventions that provide students from nondominant backgrounds with improved opportunity to use their voices and express agency to challenge contemporary inequalities within and beyond schools walls with the goal of engendering structural change.

Entry point. The Gates Foundation categorized the NSI into three entry points based on the NSI's aim statements and change ideas when they entered the initiative.[5] The entry points are (a) instructional, (b) early warning and response, and (c) well-matched postsecondary enrollment. The foundation defines the entry points as follows:

- **Instructional NSI** work with math or English-language-arts teams within schools, often including instructional coaches, special-education teachers, and English learner/multilingual teachers, to improve the quality of instruction within classrooms.
- **Early Warning and Response (EWR) NSI** work with grade-level or cross-functional teams within schools to create more supportive school environments, where young people are connected to adults, each other, and the school community.

[4] Gutiérrez, R. (2009). Helping students "play the game" and "change the game." *Todos*, 1(1).
[5] Bill & Melinda Gates Foundation personal communication, June 2021.

- **Well-Matched Postsecondary Enrollment (WMPS) NSI** work with school-based teams of counselors, service providers, district and school leaders, teachers, and other staff on evidence-based strategies and processes that support postsecondary application, enrollment, and persistence.

Inquiry cycles. Inquiry cycles are repeated, iterative tests of change conducted by network participants. Inquiry cycles may be broken into four phases—Plan-Do-Study-Act (PDSA)—that entail the following:

- Selecting a change idea and developing a plan that determines how it will be tested (Plan)
- Implementing the change idea and collecting relevant data (Do)
- Assessing the results based on the collected data (Study)
- Using the results to determine whether to adapt, abandon, or adopt the change idea (Act)

Some intermediary organizations use other conceptualizations of inquiry cycles—for example, Partners for School Innovation bases its work on Results-Oriented Cycles of Inquiry (ROCI). Cycles may have three or five phases, rather than four, or the separate phases may not be clearly defined. For more detail on how the study team defines and conceives of inquiry cycles, see Artifact Coding in Appendix C. During each cycle, outcomes are compared with predictions, and discrepancies between the two become a major source of learning.

Intermediary organization. An intermediary organization is an organization that received an NSI grant and is responsible for the facilitation and support of one or more networks and their activities. When multiple organizations work collectively to organize or support the network and its participants, we refer to the collective group as the "intermediary."

Knowledge management system. A knowledge management system is a digital platform used to organize, maintain, and share the knowledge, learning, and experiences of NSI participants.

Network. A network is a group that includes a facilitating organization and multiple schools that work together to share knowledge and practice.

Networks for School Improvement (NSI). NSI are networks funded by the Gates Foundation. NSI are composed of groups of intermediary staff and CI teams that work together to share knowledge and practice to produce more equitable student outcomes. NSI may contain subnetworks of practitioners based on a variety of factors (e.g., school district, year in which schools entered the network).

On-track threshold. A threshold set by the Gates Foundation for each outcome used to measure whether a student is on track to graduate from high school and enroll in college. For example, students earning a GPA of 3.0 or higher are considered to be on track with respect to their GPA.

Outcome area. Each intermediary organization focused its grant on improving student outcomes in one or more of the following areas:

- **Eighth or ninth grade on track.** The proportion of eighth- or ninth-grade students who meet a set of academic and behavioral outcomes related to high school graduation and college enrollment.
- **College-ready on track.** The proportion of 11th- and 12th-grade students who are on track academically to enroll in a college with a graduation rate of at least 50%.
- **Well-matched postsecondary enrollment.** The proportion of 12th-grade students who complete the steps needed to enroll in a college with a graduation rate of at least 50%.

Priority outcomes. The subset of academic and behavioral outcomes that given NSI focus on improving. The Gates Foundation focused NSI in each outcome area on improving a specific set of academic and behavioral outcomes. In practice, most NSI prioritized a subset of those outcomes rather than the full set of outcomes. For example, while the Gates Foundation expected ninth-grade on-track NSI to improve a set of five outcomes, most ninth-grade on-track NSI focused on improving three of those outcomes: GPA, course pass rates, and credit completion.

Problem of practice. A problem of practice is a current strategy or practice that network participants have identified as leading to inequitable student outcomes. Network participants address a problem of practice by conducting disciplined inquiry cycles.

Root cause. A root cause is an underlying reason for an educational challenge. Network participants identify root causes to help them understand the systems that produce inequitable outcomes for Black students, Latino students, and students experiencing poverty within their local settings.

Theory of improvement. A theory of improvement includes a set of interrelated hypotheses about how changes in certain practices or policies could lead to improved student outcomes.[6] A theory of improvement guides the work of the network and evolves as educators conduct and learn from inquiry cycles.

[6] Carnegie Foundation for the Advancement of Teaching. (2026, February 9). *Learning to improve glossary*. https://www.carnegiefoundation.org/resources/learning-to-improve-glossary

Appendix B. Initiative Characteristics

The Bill & Melinda Gates Foundation (now the Gates Foundation) funded 34 Networks for School Improvement initiatives (NSI) in total.

Exhibit B1. Intermediary Name, Network Name, Cohort, Entry Point, Outcome Area, Research Sample, and Short Name for the NSI Included in the Initiative

Intermediary name (short)	Intermediary name (full)	Network name	Cohort	Entry point	Outcome area	RQ1 analysis	RQ2 analysis	RQ3 analysis
Access ASU-2	**Access ASU**	Arizona Meta Network (AZ Meta Network)	2	Well-matched postsecondary	Well-matched postsecondary	X	X	X
AIR FNSI-1B	**American Institutes for Research**	Florida Network for School Improvement (FNSI)	1B	Instructional	College-ready on track			X
AIR Long Beach-3	**American Institutes for Research**	Long Beach Network for School Improvement (LBNSI)	3	Instructional	8th grade on track	X	X	X
Baltimore-1	**Baltimore City Public Schools**	Baltimore Secondary Literacy Improvement Community Network (BSLIC)	1	Instructional	8th grade on track; 9th grade on track	X	X	X
Baltimore-3	**Baltimore City Public Schools**	9th Grade On Track to Graduate (OTG) Improvement Network	3	Early warning and response	9th grade on track	X	X	X
Bank Street-2	**Bank Street College of Education**	Yonkers Public Schools Network for School Improvement (YPS NSI)	2	Instructional	8th grade on track			
Bank Street-3	**Bank Street College of Education**	Brooklyn South Network for School Improvement (BKS NSI)	3	Instructional	8th grade on track	X	X	
BARR-1B	**BARR Center**	BARR Network for School Improvement (BARR)	1B	Early warning and response	College-ready on track	X	X	X
City Year-2	**City Year**	City Year Network for School Improvement (City Year NSI)	2	Early warning and response	8th grade on track			X
CORE-1	**CORE**	Breakthrough Success Community (BTSC) Cohort 1	1	Early warning and response	9th grade on track	X	X	X

Intermediary name (short)	Intermediary name (full)	Network name	Cohort	Entry point	Outcome area	RQ1 analysis	RQ2 analysis	RQ3 analysis
CORE-3	**CORE**	Breakthrough Success Community (BTSC) Cohort 3	3	Early warning and response	9th grade on track	X	X	X
Denver-1B	**Denver Public Schools**	College-Ready On Track Network (NIC)	1B	Instructional	College-ready on track	X	X	X
Ed Partners-1	**California Education Partners**	On Track Improvement Collaborative	1	Early warning and response	8th grade on track; 9th grade on track			
Eskolta-2	**Eskolta School Research and Design**	Eskolta Network	2	Instructional	College-ready on track			
HTH-1	**High Tech High Graduate School of Education**	CARPE College Access Network (CARPE)	1	Well-matched postsecondary	Well-matched postsecondary	X	X	X
HTH-3	**High Tech High Graduate School of Education**	CARE Network	3	Instructional	8th grade on track	X	X	X
IFL-1	**Institute for Learning**	Dallas ISD/IFL Network for School Improvement (Dallas ISD/IFL NSI)	1	Instructional	8th grade on track; 9th grade on track			
KIPP-2	**KIPP Foundation**	Academics and Counseling Excellence Network (ACE Network)	2	Well-matched postsecondary	College-ready on track; well-matched postsecondary			
NCS-1	**UChicago Network for College Success**	Chicago School Partner Network or Network for College Success (NCS)	1	Early warning and response	9th grade on track	X	X	X
NCS-3	**UChicago Network for College Success**	Freshman Success for Equity Improvement Network (FS4EIN)	3	Early warning and response	9th grade on track	X	X	X
New Visions-1	**New Visions for Public Schools**	College Readiness Network for School Improvement (CR NSI)	1	Early warning and response	9th grade on track	X	X	X
New Visions-3	**New Visions for Public Schools**	Instructional Network for School Improvement (INSI)	3	Instructional	9th grade on track	X	X	X

Intermediary name (short)	Intermediary name (full)	Network name	Cohort	Entry point	Outcome area	RQ1 analysis	RQ2 analysis	RQ3 analysis
NTN-1B	**New Tech Network**	NTN College Access Network	1B	Well-matched postsecondary	Well-matched postsecondary	X	X	X
NYC DOE-2	**New York City Department of Education**	Networked Improvement Community for Multilingual Learners (CL ML NIC)	2	Instructional	8th grade on track			X
Partners ESA-2	**Partners in School Innovation**	East Side Alliance Transformation Network (ESA Transformation Network)	2	Instructional	8th grade on track	X	X	
Partners MGIT-2	**Partners in School Innovation**	Middle Grade Improvement Team Network	2	Instructional	8th grade on track	X	X	X
Partners MGS1-2	**Partners in School Innovation**	Middle Grades Success Network	2	Instructional	8th grade on track	X	X	X
Partners OTSN-2	**Partners in School Innovation**	On Track for Success Network	2	Instructional	8th grade on track	X	X	X
Promise-1B	**The Commit Partnership**	Promise Network for School Improvement (Promise Network)	1B	Well-matched postsecondary	Well-matched postsecondary	X	X	X
RISE-2	**Connecticut RISE Network**	Connecticut RISE Network (RISE or RISE Network)	2	Early warning and response	9th grade on track	X	X	
Teaching Matters-3	**Teaching Matters**	Teaching Matters Network for School Improvement (Teaching Matters NSI)	3	Instructional	8th grade on track	X	X	X
Teach Plus-3	**Teach Plus**	Teacher-Led Network for School Improvement in Chicago	3	Instructional	8th grade on track	X	X	X
Tulare County-3	**Tulare County Office of Education**	Central Valley Networked Improvement Community: College-Ready (CVNIC: College-Ready)	3	Instructional	College-ready on track	X	X	X
TxNSI-2	**Texas Network for School Improvement**	Texas Network for School Improvement (TXNSI)	2	Instructional	8th grade on track			

Note. RQ = research question.

Exhibit B2. Definitions of NSI Outcomes

Outcome	Description	Threshold for a student to be on track
8th grade on track		
Grade point average (GPA) for core courses (math, English language arts [ELA], science, and social studies)	8th-grade GPA based on core courses (4-point scale)	GPA for core courses of at least 3.0
Math and ELA course grades	The proportion of math and ELA courses for which students earned at least a C	Received no Ds or Fs in math and ELA courses
Math test scores	Score on the state standardized math assessment	Scoring at least at the proficient level on the state math assessment and earning at least a 3.0 GPA in math
ELA test scores	Score on the state standardized ELA assessment	Scoring at least at the proficient level on the state ELA assessment and earning at least a 3.0 GPA in ELA
Attendance rate	Percentage of days a student attended school	Attended at least 96% of school days
Received no suspensions	Whether a student received no out-of-school suspensions	Received no out-of-school suspensions
9th grade on track		
GPA for all courses	9th-grade GPA based on all courses (4 point scale)	GPA for all courses at least 3.0
Share of core courses passed	The proportion of core courses for which a student earned at least a C	One or fewer course failures in the core subject areas
Earned at least 5 credits	Whether a student earned at least 5 credits	Earned at least 5 credits
Attendance rate	Percentage of days a student attended school	Attended at least 96% of school days
Received no suspensions	Whether a student received no out-of-school suspensions	Received no out-of-school suspensions
College-ready on-track		
High school advanced course taking	Whether an 11th- or 12th-grade student completes at least one advanced course (Advanced Placement, International Baccalaureate, or dual enrollment)	N/A
High school math proficiency	Whether an 11th- or 12th-grade student scored at the proficient level on a high school state assessment in math	
High school ELA proficiency	Whether an 11th- or 12th-grade student scored at the proficient level on a high school state assessment in ELA	
High school grade point average	Whether an 11th- or 12th-grade student earned a GPA of 3.0 or higher.	
On-time high school graduation	Whether a student graduated from high school in 4 years with a regular high school diploma	

Outcome	Description	Threshold for a student to be on track
Postsecondary enrollment rate	Whether a 12th-grade student enrolled in any postsecondary institution in the fall following their 12th-grade year	
Postsecondary enrollment rate for institutions with a graduation rate of at least 50%	Whether a 12th-grade student enrolled in a postsecondary institution with a graduation rate at least 50% in the fall following their 12th-grade year	
Well-matched postsecondary enrollment		
Exam rate	Whether a 12th-grade student took the SAT and/or ACT	N/A
FAFSA completion rate	Whether a 12th-grade student completed a FAFSA form by December of the school year	
Secured postsecondary plan	Whether a 12th-grade student completed the FAFSA, submitted applications to at least three colleges, and completed at least one college entrance exam.	
Postsecondary enrollment rate	Whether a 12th-grade student enrolled in any postsecondary institution in the fall following their 12th-grade year	
Postsecondary enrollment rate for institutions with a graduation rate of at least 50%	Whether a 12th-grade student enrolled in a postsecondary institution with a graduation rate at least 50% in the fall following their 12th-grade year	

Note. The initiative did not define on-track thresholds for the college-ready on-track and well-matched postsecondary NSI. FAFSA = Free Application for Federal Student Aid.

Appendix C. Evaluation Sample Characteristics

The study sample includes schools associated with 25 Networks for School Improvement (NSI). We collected data from all cohorts in 2020–21. Beginning in 2021–22, the study team collected data for NSI in Grant Years 3 or 5. Exhibit C1 shows the number of NSI included in the evaluation and the timing of the grants by cohort. In some NSI, all schools began participating in the first year of the grant; in other NSI, some schools began participating in later years. This appendix details the data collection activities by data source in each school year.

Exhibit C1. Number of NSI and Year of Grant, by Cohort

Cohort	Number of NSI	2018–19	2019–20	2020–21	2021–22	2022–23	2023–24	2024–25
1	5	1	2	3	4	5		
1B/2	10		1	2	3	4	5	
3	10			1	2	3	4	5

Note. Data for the years shaded in green and yellow serve as the focus of this report. Most of the descriptive results presented are based on Grant Year 5 for each Cohort: 2022–23 data for Cohort 1, 2023-24 data for Cohort 2, and 2024–25 data for Cohort 3. Regression models include additional years of data shaded in yellow.

Exhibit C2. Number of NSI in Evaluation, by Cohort

Cohort	NSI launch year	Number of NSI
Cohort 1	2018–19	5
Cohort 1B/2	2019–20	10
Cohort 3	2020–21	10

Exhibit C3. Number of NSI in Evaluation, by Entry Point

Entry point	Number of NSI
Instructional	13
Early warning and response	8
Well-matched postsecondary	4

Exhibit C4. Number of NSI in Evaluation, by Outcome Area

Outcome Area	Number of NSI
8th grade on track	10
9th grade on track	9
College-ready on track	3
Well-matched postsecondary enrollment	4

Note. Baltimore-1 is present in both 8th grade on track and 9th grade on track.

Exhibit C5. Participation in Evaluation Data Collection, by NSI and School Year

NSI	Intermediary interviews					Continuous improvement artifacts					NSI school leader surveys					Case study interviews				
	2020–21	2021–22	2022–23	2023–24	2024–25	2020–21	2021–22	2022–23	2023–24	2024–25	2020–21	2021–22	2022–23	2023–24	2024–25	2020–21	2021–22	2022–23	2023–24	2024–25
Access ASU-2	●	●		●		●	●				●	●		●		●				
AIR Long Beach-3	●		●			●		●		●	●		●		●					
Baltimore-1	●		●			●		●			●		●		●	●	●	●		●
Baltimore-3	●		●		●	●		●			●		●		●	●				
Bank Street-3	●		●		●	●		●			●		●		●	●				
BARR-1B	●	●		●		●	●				●	●		●	●					
CORE-1	●		●			●		●			●		●		●					
CORE-3a	●	●	●		●		●	●				●	●	●	●					
Denver-1B	●	●		●		●	●				●	●		●						
HTH-1	●		●			●		●			●		●		●					
HTH-3	●		●		●	●		●			●		●		●					
NCS-1	●		●			●		●			●		●		●					
NCS-3	●		●		●	●		●			●		●		●					
New Visions-1	●		●			●		●			●		●		●					
New Visions-3[b]	●	●	●		●		●	●				●	●		●	●	●	●	●	
NTN-1B	●	●		●		●	●				●	●		●	●	●	●	●	●	●
Partners ESA-2	●	●		●		●	●				●	●		●	●					

NSI	Intermediary interviews					Continuous improvement artifacts					NSI school leader surveys					Case study interviews				
	2020–21	2021–22	2022–23	2023–24	2024–25	2020–21	2021–22	2022–23	2023–24	2024–25	2020–21	2021–22	2022–23	2023–24	2024–25	2020–21	2021–22	2022–23	2023–24	2024–25
Partners MGIT-2[c]		●		●			●					●		●	●					
Partners MGS1-2	●	●		●		●	●				●	●		●	●					
Partners OTSN-2	●	●		●		●	●				●	●		●	●					
Promise-1B	●	●		●		●	●				●	●		●	●					
RISE-2	●	●		●		●	●				●	●		●	●	●	●	●	●	●
Teach Plus-3	●		●		●	●		●			●		●		●					
Teaching Matters-3	●		●		●	●		●			●		●		●					
Tulare County-3[b]	●	●	●		●		●	●				●	●		●		●	●	●	●

[a] Schools in CORE-3 did not begin participating until after the 2020–21 school year, so we collected data for them in the 2021–22 school year.

[b] Schools in New Visions-3 and Tulare County-3 started midway through the 2020–21 school year, so we collected school-level data from them in the 2021–22 school year.

[c] Schools in Partners MGIT-2 did not begin participating until January 2022, so we did not collect data from them until the 2021–22 school year.

Exhibit C6. Schools and Districts in the Evaluation Sample, by NSI and School Year

NSI	Number of schools served in 2020–21	Number of schools served in 2021–22	Number of schools served in 2022–23	Number of schools served in 2023–24	Number of schools served in 2024–25	Unduplicated number of schools served from 2020–21 to 2024–25	Unduplicated number of districts served from 2020–21 to 2024–25
Access ASU-2	24	24	24	24	24	24	3
AIR Long Beach-3	10	10	10	19	19	21	1
Baltimore-1	25	25	25	8	7	26	1
Baltimore-3	12	12	12	23	23	25	1
Bank Street-3	11	11	11	21	21	21	2
BARR-1B	31	31	31	31	31	33	22
CORE-1	28	24	23	20	20	28	9
CORE-3	0	11	11	23	23	23	5
Denver-1B	10	22	23	21	21	32	1
HTH-1	19	31	27	23	23	40	21
HTH-3	16	16	14	22	24	35	13
NCS-1	18	18	13	15	15	23	1
NCS-3	11	11	11	11	20	20	1
New Visions-1	43	43	62	0	0	63	31
New Visions-3	18	16	14	23	23	27	15
NTN-1B	25	49	49	49	49	49	15
Partners ESA-2	8	7	9	9	9	10	5
Partners MGI-2	0	8	8	8	8	8	1
Partners MGS1-2	10	8	7	0	0	10	1
Partners OTSN-2	13	13	13	12	12	14	1

Promise-1B	57	57	57	57	57	57	11
RISE-2	10	9	9	9	9	10	9
Teach Plus-3	15	15	14	43	43	47	1
Teaching Matters-3	16	16	13	14	14	19	12
Tulare County-3	0	14	14	15	15	17	11

Exhibit C7. Number of Schools That Provided Data and Response Rates in 2020–2021, by NSI

NSI	Number of schools served in 2020–21	Continuous improvement artifacts				NSI school leader survey	
		Number of schools that provided Phase 2	Response rate	Number of schools that provided Phase 3	Response rate	Number of schools that provided data	Response rate
Access ASU-2	24	16	67%	24	100%	13	54%
AIR Long Beach-3	10	5	50%	10	100%	10	100%
Baltimore-1[a,b]	25	0	0%	25	100%	12	48%
Baltimore-3[b]	12	11	92%	12	100%	3	25%
Bank Street-3[a]	11	0	0%	11	100%	7	64%
BARR-1B	31	9	29%	27	87%	21	68%
CORE-1	28	12	43%	27	96%	23	82%
CORE-3[c]							
Denver-1B	10	8	80%	10	100%	4	40%
HTH-1	18	18	100%	18	100%	9	50%
HTH-3	16	12	75%	15	94%	9	56%
NCS-1	18	12	67%	18	100%	10	56%
NCS-3[ab]	11	0	0%	11	100%	4	36%
New Visions-1	43	9	21%	9	21%	26	60%
New Visions-3[d]	18						
NTN-1B[e]	25	1	4%	25	100%	18	72%
Partners ESA-2	8	7	88%	8	100%	8	100%
Partners MGIT-2[f]							
Partners MGS1-2	10	8	80%	10	100%	1	10%
Partners OTSN-2[a]	13	0	0%	13	100%	9	69%
Promise-1B[b]	57	35	61%	55	96%	9	16%
RISE-2[e]	10	2	20%	10	100%	7	70%
Teach Plus-3[g]	15	0	0%	0	0%	12	80%
Teaching Matters-3	16	6	38%	12	75%	12	75%
Tulare County-3[d]							

Note. See Appendix D for a discussion of Phase 2 and Phase 3 artifacts.

[a] Although we received artifacts from Baltimore-1, Bank Street-3, NCS-3, and Partners OTSN-2, coders determined that none were related to individual inquiry cycles conducted by school teams. So, we report zero schools provided Phase 2 artifacts for these NSI.

[b] We did not include school leader survey responses from Baltimore-1, Baltimore-3, NCS-3, Partners MGS1–2, or Promise-1B in our analyses because they did not achieve specified response rates.

[c] Schools in CORE-3 did not begin participating until after the 2020–21 school year, so we collected data for them in 2021–22 school year.

[d] Schools in New Visions-3 and Tulare County-3 started midway through the 2020–21 school year, so we collected school-level data for them in the 2021–22 school year.

[e] Although we received artifacts from NTN-1B and RISE-2, coders determined that artifacts showed evidence of individual inquiry cycles in fewer than three schools. To protect the anonymity of participants, we did not include the data in the report.

[f] Schools in Partners MGIT-2 did not begin participating until January 2022, so we did not collect data from them until the 2021–22 school year.

[g] Although we received artifacts from Teach Plus-3, coders determined that none of the artifacts were related to continuous improvement work conducted by school teams. So, we report that zero schools provided Phase 2 or Phase 3 artifacts.

Exhibit C8. Number of Schools That Provided Data and Response Rates in 2021–22, by NSI

NSI	Number of schools served 2021–22	Continuous improvement artifacts				NSI school leader survey	
		Number of schools that provided Phase 2	Response rate	Number of schools that provided Phase 3	Response rate	Number of schools that provided data	Response rate
Access ASU-2	24	24	100%	24	100%	19	79%
AIR-Long Beach-3	10						
Baltimore-1	25						
Baltimore-3	12						
Bank Street-3	11						
BARR-1B	31	18	58%	28	90%	25	81%
CORE-1	24						
CORE-3[a]	11	2	18%	11	100%	8	73%
Denver-1B	22	16	73%	22	100%	19	86%
HTH-1	31						
HTH-3	16						
NCS-1	18						
NCS-3	11						
New Visions-1	43						
New Visions-3[a]	16	9	63%	16	100%	10	63%
NTN-1B	49	45	86%	49	100%	37	76%
Partners ESA-2	7	7	100%	7	100%	7	100%
Partners MGIT-2[b]	8	0	0%	8	100%	4	50%
Partners MGS1-2	8	4	50%	8	100%	5	63%
Partners OTSN-2	13	2	15%	13	100%	7	54%
Promise-1B[b]	57	0	0%	43	75%	23	40%
RISE-2	9	7	78%	9	100%	7	78%

NSI	Number of schools served 2021–22	Continuous improvement artifacts				NSI school leader survey	
		Number of schools that provided Phase 2	Response rate	Number of schools that provided Phase 3	Response rate	Number of schools that provided data	Response rate
Teach Plus-3	15						
Teaching Matters-3	16						
Tulare County-3[a]	14	13	93%	14	100%	11	79%

Note. See Appendix D for a discussion of Phase 2 and Phase 3 artifacts.

[a] We collected and analyzed data from CORE-3, New Visions-3, and Tulare County-3 in 2021–22 when they were in their second grant year because we did not collect data from them in 2020–21.

[b] Although we received artifacts from Partners MGIT-2 and Promise-1B, coders determined that none of the artifacts were related to individual inquiry cycles conducted by school teams. So, we report that zero schools provided Phase 2 artifacts.

Exhibit C9. Number of Schools That Provided Data and Response Rates in 2022–23, by NSI

NSI	Number of schools served 2022–23	Continuous improvement artifacts				NSI school leader survey	
		Number of schools that provided Phase 2	Response rate	Number of schools that provided Phase 3	Response rate	Number of schools that provided data	Response rate
Access ASU-2	24						
AIR Long Beach-3	10	10	100%	10	100%	8	80%
Baltimore-1	25	11	44%	25	100%	16	64%
Baltimore-3	12	10	83%	12	100%	9	75%
Bank Street-3[a]	11	0	0%	11	100%	8	73%
BARR-1B	31						
CORE-1	23	9	39%	23	100%	16	70%
CORE-3	11	4	36%	6	55%	11	100%
Denver-1B	19						
HTH-1	27	25	93%	27	100%	24	89%
HTH-3	13	9	69%	13	100%	11	85%
NCS-1	13	8	62%	13	100%	13	100%
NCS-3	11	5	45%	11	100%	11	100%
New Visions-1[a]	9	0	0%	9	100%	9	100%
New Visions-3	14	13	93%	14	100%	12	86%
NTN-1B	49						
Partners ESA-2	9						
Partners MGIT-2	8						
Partners MGS1-2	7						
Partners OTSN-2	14						
Promise-1B	57						
RISE-2	9						
Teach Plus-3	14	14	100%	14	100%	13	93%
Teaching Matters-3	16	12	75%	16	100%	12	75%
Tulare County-3	14	13	93%	14	100%	11	79%

Note. See Appendix D for a discussion of Phase 2 and Phase 3 artifacts.

[a] Although we received artifacts from Bank Street-3 and New Visions-1, coders determined that none of the artifacts were related to individual inquiry cycles conducted by school teams. So, we report that zero schools provided Phase 2 artifacts.

Exhibit C10. Number of Schools That Provided Data and Response Rates in 2023–24, by NSI

NSI	Number of schools served 2023–24	Continuous improvement artifacts				NSI school leader survey	
		Number of schools that provided Phase 2	Response rate	Number of schools that provided Phase 3	Response rate	Number of schools that provided data	Response rate
Access ASU-2	24	9	38%	24	100%	14	58%
AIR Long Beach-3							
Baltimore-1							
Baltimore-3							
Bank Street-3							
BARR-1B	31	26	84%	30	97%	22	71%
CORE-1							
CORE-3							
Denver-1B	22	20	91%	21	95%	12	55%
HTH-1							
HTH-3							
NCS-1							
NCS-3							
New Visions-1							
New Visions-3							
NTN-1B	49	44	90%	49	100%	33	67%
Partners ESA-2	9	9	100%	9	100%	7	78%
Partners MGIT-2	8	3	38%	7	88%	5	62%
Partners MGS1-2							
Partners OTSN-2	12	0	0%	12	100%	6	50%
Promise-1B	57	0	0%	57	100%	30	53%
RISE-2	9	9	100%	9	100%	8	89%
Teach Plus-3							
Teaching Matters-3							
Tulare County-3							

Exhibit C11. Number of Schools That Provided Data and Response Rates in 2024–25, by NSI

NSI	Number of schools served 2024–25	Continuous improvement artifacts				NSI school leader survey	
		Number of schools that provided Phase 2	Response rate	Number of schools that provided Phase 3	Response rate	Number of schools that provided data	Response rate
Access ASU-2							
AIR Long Beach-3	17	17	100%	17	100%	8	47%
Baltimore-1							
Baltimore-3	21	14	67%	14	67%	16	76%
Bank Street-3	21	13	62%	21	100%	14	67%
BARR-1B							
CORE-1							
CORE-3	23	9	39%	23	100%	10	43%
Denver-1B							
HTH-1							
HTH-3	21	8	38%	20	95%	16	76%
NCS-1							
NCS-3	20	16	80%	19	95%	14	70%
New Visions-1							
New Visions-3	23	22	96%	23	100%	15	65%
NTN-1B							
Partners ESA-2							
Partners MGIT-2							
Partners MGS1-2							
Partners OTSN-2							
Promise-1B							
RISE-2							
Teach Plus-3	25	21	84%	25	100%	19	76%
Teaching Matters-3	15	15	100%	15	100%	9	60%
Tulare County-3	11	10	91%	11	100%	9	82%

Exhibit C12. Number of Schools With Climate Data, by NSI and Grant Year

NSI	Number of schools served across all years	Grant Year 1		Grant Year 3		Grant Year 5	
		n of schools with climate data	Response rate	*n* of schools with climate data	Response rate	*n* of schools with climate data	Response rate
AIR-Long Beach-3	9	9	100%	0	0%	0	0%
Bank Street-3	10	10	100%	10	100%	0	0%
CORE-1[a]	7	0	0%	7	100%	0	0%
NCS–1	18	0	0%	17	94%	13	72%
NCS–3	11	10	91%	10	91%	0	0%
New Visions-1	9	0	0%	9	100%	9	100%
New Visions-3	16	16	100%	14	88%	0	0%
Partners MGS1–2	10	10	100%	8	80%	0	0%
Partners OTSN–2	14	13	93%	13	93%	12	86%
Teach Plus-3	13	0	0%	13	100%	0	0%
Teaching Matters-3	13	11	85%	13	100%	0	0%

[a] For CORE-1, the number of schools served includes only schools in Long Beach Unified and Fresno Unified (districts for which we have climate data).

Exhibit C13. Number of Schools With Sustainability Survey Data, by NSI

NSI	Number of schools invited to complete survey	Number of completions	Response Rate
Baltimore-1	27	7	26%
BARR-1B	31	11	35%
CORE–1	23	3	13%
RISE–2	9	4	44%
HTH–1	36	9	25%
NCS-1	20	7	35%
New Visions-1	62	17	27%
NTN-1B	49	16	33%
Partners ESA-2	9	5	56%
Partners MGI-2	8	4	50%
Partners MGS1-2	7	4	57%
Partners OTSN-2	13	2	15%
Promise-1B	57	6	11%

Note. The sustainability survey was sent to school leaders in networks that were no longer receiving direct foundation support as of spring 2025.

Exhibit C14. Characteristics of Schools in the NSI Evaluation

NSI	Number of schools	Majority school type	Majority region	Median total school enrollment	Percentage of schools by grade level[a]			Percentage of schools by locale			
					Elementary[c]	Middle[c]	High[c]	Urban	Suburban	Town	Rural
Access ASU-2	24	Regular	West	2,217	0%	0%	100%	75%	25%	0%	0%
AIR Long Beach-3	22	Regular	West	866	0%	95%	5%	86%	9%	5%	0%
Baltimore-1	27	Regular	South	481	0%	74%	26%	100%	0%	0%	0%
Baltimore-3	25	Regular	South	715	0%	0%	100%	100%	0%	0%	0%
Bank Street-3	21	Regular	Northeast	322	0%	100%	0%	100%	0%	0%	0%
BARR-1B	33	Regular	Midwest	1,644	0%	0%	100%	24%	64%	6%	6%
CORE-1	31	Regular	West	1,941	0%	0%	100%	87%	10%	3%	0%
CORE-3	23	Regular	West	1,683	0%	0%	100%	74%	26%	0%	0%
Denver-1B	32	Regular	West	272	0%	0%	100%	91%	0%	0%	6%
HTH-1	41	Regular	West	1,796	0%	0%	100%	46%	49%	0%	5%
HTH-3	38	Regular	West	649	0%	95%	5%	58%	39%	0%	3%
NCS-1	26	Regular	Midwest	818	0%	0%	100%	100%	0%	0%	0%
NCS-3	20	Regular	Midwest	387	0%	0%	100%	100%	0%	0%	0%
New Visions-1	63	Regular	Northeast	410	0%	0%	100%	100%	0%	0%	0%
New Visions-3	43	Regular	Northeast	415	0%	0%	100%	100%	0%	0%	0%
NTN-1B	49	Regular	South	1,225	0%	0%	100%	47%	27%	6%	20%
Partners ESA-2	10	Regular	West	551	0%	100%	0%	100%	0%	0%	0%
Partners MGI-2	8	Regular	West	493	0%	88%	12%	100%	0%	0%	0%
Partners MGS1-2	11	Regular	Northeast	394	0%	100%	0%	100%	0%	0%	0%
Partners OTSN-2	14	Regular	Northeast	415	0%	100%	0%	100%	0%	0%	0%
Promise-1B	57	Regular	South	1,899	0%	0%	100%	65%	33%	0%	2%
RISE-2	10	Regular	Northeast	1,274	0%	0%	100%	60%	40%	0%	0%

NSI	Number of schools	Majority school type	Majority region	Median total school enrollment	Percentage of schools by grade level[a]			Percentage of schools by locale			
					Elementary[c]	Middle[c]	High[c]	Urban	Suburban	Town	Rural
Teach Plus-3	47	Regular	Midwest	327	0%	100%	0%	100%	0%	0%	0%
Teaching Matters-3	30	Regular	Northeast	312	0%	87%	13%	100%	0%	0%	0%
Tulare County-3	17	Regular	West	1,264	0%	0%	100%	24%	18%	29%	29%

Source. NSI school list maintained by Mathematica Policy Research. Numbers are based on all schools participating in NSI in 2020–21, 2021–22, 2022–23, 2023-24, or 2024–25. Data are based on the Common Core of Data, 2020–21.

Note. The *n* for the summary information in this exhibit is lower than the *n* of schools in the analysis because some schools are too new to have data in the National Center for Education Statistics Common Core of Data, the source of information for the NSI school list maintained by Mathematica Policy Research.

[a] Not all grade-level categories by NSI sum to 100% because some schools' grade levels were classified as "other," which is not shown.

[c] Grade levels are categorized using the National Center for Education Statistics definition of instructional levels. The four instructional levels are elementary (lowest grade of prekindergarten up to Grade 3; highest grade up to Grade 8), middle (lowest Grades 4–7; highest grade up to Grades 4–9), high (lowest Grade 7 up to Grade 12; highest Grade 12), and other (all other configurations, including prekindergarten, kindergarten, or Grades 1–12).

Exhibit C15. Demographic Characteristics of Schools in NSI Evaluation

NSI	Number of schools	Average percentage of students				
		Eligible for free or reduced-price lunch	Black	Latino or Hispanic	White	Asian
Access ASU-2	24		8%	71%	14%	2%
AIR Long Beach-3	22	64%	12%	59%	13%	9%
Baltimore-1	27	71%	75%	17%	6%	1%
Baltimore-3	25	67%	81%	11%	6%	1%
Bank Street-3	21	85%	78%	14%	3%	2%
BARR-1B	33	59%	19%	46%	27%	3%
CORE-1	31	78%	13%	66%	9%	9%
CORE-3	23	84%	4%	73%	6%	16%
Denver-1B	32	72%	16%	64%	14%	2%
HTH-1	41	68%	7%	73%	11%	5%
HTH-3	38	74%	7%	69%	11%	7%
NCS-1	26		45%	47%	4%	2%
NCS-3	20		46%	44%	6%	3%
New Visions-1	63	83%	38%	48%	6%	6%
New Visions-3	43	86%	34%	51%	4%	8%
NTN-1B	49	69%	6%	78%	13%	1%
Partners ESA-2	10	46%	2%	59%	6%	28%
Partners MGI-2	8	96%	27%	69%	2%	1%
Partners MGS1-2	11	99%	59%	27%	7%	3%
Partners OTSN-2	14	99%	43%	19%	20%	11%
Promise-1B	57	77%	25%	64%	6%	3%
RISE-2	10	63%	21%	45%	26%	4%
Teach Plus-3	47		64%	33%	1%	0%
Teaching Matters-3	30	92%	31%	61%	3%	2%
Tulare County-3	17	74%	1%	78%	15%	3%

Source. NSI school list maintained by Mathematica Policy Research. Numbers are based on all schools participating in an NSI in 2020–21, 2021–22, 2022–23, 2023-24, or 2024-25. Data are based on the Common Core of Data, 2020–21.
Note. The *n* for the summary information in this exhibit is lower than the *n* of schools in the analysis because some schools are too new to have data in the National Center for Education Statistics Common Core of Data.
Free or reduced-price lunch data were not available for schools in Access ASU-2, NCS-1, NCS-3, or Teach Plus-3.

Appendix D. Description of Data Collection and Response Rates

This appendix describes the data collection process and response rates for each of the data sources used in the study.

Continuous Improvement (CI) Artifacts

The study team gathered artifacts that CI teams generated during their work, along with resources provided by intermediaries to support CI, to learn about the implementation of CI in Networks for School Improvement (NSI) schools. An artifact is any documentation created by the intermediary, NSI, a school, or a CI team at any phase of the CI process. For example, artifacts include training materials provided by an NSI, root cause analyses and descriptions of the interventions tested (created by an intermediary or school), CI team meeting notes, CI templates created by an intermediary, documents created during inquiry cycles, internal reports, and data collections or visualizations.

Artifact Data Collection

The team collected artifacts for the 2020–21, 2021–22, 2022–23, 2023–24, and 2024–25 school years. For the 2020–21 school year, we collected artifacts separately for the fall and spring. Beginning in late January 2021, the study team contacted representatives from each NSI in Cohorts 1, 1B, and 2. We requested all artifacts created by these intermediaries, as well as the districts, schools, or CI teams they supported, that documented CI processes at the network, district, and school levels in the fall. To better define this request, we created a short document describing examples of possible artifacts (Exhibit D1); however, we did not limit our request to only the artifact types listed.

Exhibit D1. Possible Artifact Types

Intermediary plans and tools	Training materials	School CI records
• Intermediary data collection plan • Template(s) used to collect data from districts and schools • Tools to monitor school-level CI progress • Within-grantee evaluation or progress reports (e.g., internal monitoring or evaluation) • Other resources provided by an intermediary to guide the CI process	• Training schedules • Agendas • Attendance lists • Frameworks used • Activity rubrics and templates • Products created from CI trainings • Other resources developed by an intermediary to support the CI process	• CI meeting schedules • CI meeting agendas • CI meeting attendance lists • Meeting notes and action items • Materials generated during inquiry cycles • Evaluation or progress reports • Other materials created by a network school during the CI process

Note. This exhibit was included in the fall 2020 Artifact Collection Overview handout. CI = continuous improvement.

We asked each intermediary to upload all artifacts to our secure project website or to provide the study team with the locations in which the artifacts were stored so that we could download them. To collect artifacts for the remainder of the school year, the team contacted these NSI again in July 2021.

Because the initial artifact collection contained many network-level documents without reference to the actions of school teams, and we found such documents to be of limited usefulness in our initial analyses, we subsequently asked intermediaries to provide only artifacts that related specifically to CI activity at the school level. For the 2021–22 school year, we also requested artifacts twice. In November 2021, we asked for artifacts related to up to two CI cycles. Then, in June 2022, we asked for any remaining artifacts.

For the 2022–23, 2023-24, and 2024-25 school years, we requested artifacts for the full school year once, in May, at the end of each school year.[7]

Exhibit D2 shows the number of artifacts coded from 2020–21 to 2024–25 for each NSI.

Exhibit D2. Number of Artifacts Coded From 2020–21 to 2024–25, by NSI

NSI	Number of artifacts coded for 2020–21 a	Number of artifacts coded for 2021–22 a	Number of artifacts coded for 2022–23	Number of artifacts coded for 2023-24	Number of artifacts coded for 2024–25
Access ASU-2	149	34	n/a	183	n/a
AIR Long Beach-3	24	n/a	336	n/a	194
Baltimore-1	124	n/a	632	n/a	n/a
Baltimore-3	318	n/a	388	n/a	1,444
Bank Street-3	56	n/a	28	n/a	78
BARR-1B	87	106	n/a	127	n/a
CORE-1	57	n/a	29	n/a	n/a
CORE-3	n/a	17	87	n/a	40
Denver-1B	163	190	n/a	70	n/a
HTH-1	281	n/a	312	n/a	n/a
HTH-3	142	n/a	255	n/a	59
NCS-1	83	n/a	43	n/a	n/a
NCS-3	n/a	n/a	36	n/a	49
NTN-1B	388	594	n/a	101	n/a
New Visions-1	92	n/a	7	n/a	n/a
New Visions-3	27	117	129	n/a	371
Partners ESA-2	194	177	n/a	39	n/a
Partners MGIT-2	131	40	n/a	18	n/a
Partners MGS1-2	134	51	n/a	n/a	n/a

[7] We initially collected artifact data twice a year, so we could report results on fall implementation to NSI relatively quickly. But it proved more efficient and less burdensome to collect artifacts once a year.

NSI	Number of artifacts coded for 2020–21 a	Number of artifacts coded for 2021–22 a	Number of artifacts coded for 2022–23	Number of artifacts coded for 2023-24	Number of artifacts coded for 2024–25
Partners OTSN-2	n/a	26	n/a	23	n/a
Promise-1B	105	10	n/a	29	n/a
RISE-2	54	35	n/a	15	n/a
Teach Plus-3	n/a	n/a	9	n/a	185
Teaching Matters-3	41	n/a	48	n/a	248
Tulare County-3	96	72	32	n/a	151
Total	**2597**	**1435**	2,371	605	2,819

[a] The number of artifacts coded for 2020–21 and 2021–22 will be included in an updated version of this report.
Note. n/a = NSI was not in the sample for that year.

Artifact Coding

Artifacts were coded by a team of research staff (referred to as "coders" or "the coding team"). The coders were provided with a copy of the coding protocol and a 90-minute training in its use. Coders were then certified, based on coding a predetermined set of artifacts. Their results were compared with a set of codes created by team leadership. To be certified, coders needed to meet a 70% threshold of agreement with the leadership-created codes.

Artifact coding took place in three phases:

- Phase 1: Artifact description
- Phase 2: Inquiry cycles
- Phase 3: Core parameters of CI

In the artifact description phase, each individual artifact was coded for basic descriptive information, including the intermediary, network, school, and CI team[8] to which the artifact pertained; the type of artifact; and the aspects of CI that the artifact described. We divided the aspects of CI into two categories: inquiry cycles (Phase 2) and core parameters (Phase 3). The inquiry cycle category was further divided into plan-do-study-act (PDSA), and the core parameters category was divided into features of CI other than cycles (e.g., root cause analysis).

The primary purpose of the artifact description phase was to sort artifacts into the groups used for inquiry cycle and core parameter coding. For example, all artifacts related to the plan phase of the first inquiry cycle conducted by a CI team were grouped together.

[8] Baltimore-1 did not have CI teams. Instead, each school had one or more fellows who worked on CI cycles independently. For the purposes of artifact coding, we considered each of these fellows to be a distinct team.

Most schools had only one CI team, but about 6% of schools had two or more teams. For inquiry cycle coding (Phase 2), each team's cycles were coded separately. For core parameter coding (Phase 3), we coded one team per school. If a school had more than one team, coders chose which team to code based on the following criteria. If only one team had evidence of cycles, coders completed Phase 3 coding for that team. If more than one team had evidence of cycles, or no teams showed evidence of cycles, coders selected the team with the most detailed evidence of CI activities for Phase 3 coding.

In the inquiry cycle and core parameter phases, the coders determined whether the artifacts as a group contained clear evidence of specific features or aspects of CI. The absence of such evidence does not necessarily imply that the feature or aspect did not occur in the team's CI activities—it simply indicated that the feature or aspect was not recorded in the artifacts.

The study team understands that not every NSI follows the PDSA model for inquiry cycles. To compare inquiry cycles across NSI, the coding team used the following definitions to sort artifacts into cycle phases, even if the specific language describing the phase differed:

- Plan: Developing a plan that identifies an intervention and determines how that intervention will be tested
- Do: Implementing the plan
- Study: Assessing the results
- Act: Reflecting on the results and adjusting the plan by adopting the intervention, adapting it, or abandoning it

If CI team artifacts documented activities related to the above inquiry cycle tasks, they were coded according to their respective phases, whether or not the CI teams divided their processes into these same four phases or explicitly named the phases of their inquiry cycles in the same way. A single artifact might have been coded with respect to multiple phases, depending on the information it contained.

Artifacts for the 2020–21 school year were coded in two batches: fall 2020 artifacts between April and July 2021, and spring 2021 artifacts between September 2021 and March 2022. When this coding was complete, we refined and updated the coding protocol, although it followed the same basic structure. The changes primarily improved skip logic, provided more detailed answer options, and incorporated the equity framework we had developed with the Research Question (RQ)1 and RQ3 teams. Artifacts for the 2021–22 school year were coded between April 2022 and February 2023, and artifacts for the 2022–23, 2023–24, 2024–25 schools years were coded between June and September going into the next school year (e.g., 2024–25 artifacts were coded between June and September 2025). We counted an artifact as coded if it was provided by the intermediary and was determined to be relevant to the coding during the artifact description phase of the coding process.

Response Rate

For response rates by NSI, see Exhibits C7, C8, C9, C10, and C11.

NSI School Leader Survey

NSI School Leader Survey Design

The study team designed a web-based NSI school leader survey to gather information on the school context in which the initiative was implemented. Although the survey was designed with school principals in mind, questions were worded so that any school-based administrator (e.g., assistant principal, dean, teacher leader) could respond. Most items were closed-ended.

The survey contains four or five sections, depending on the year of administration:

1. Schools' response to the COVID-19 pandemic (School Years 2021 and 2022 only)
2. Schools' approaches to equity
3. Time and opportunities provided to educators to plan and collaborate with one another
4. Participation in the NSI
5. Leaders' professional background and their schools' characteristics

Retired school administrators who work at the American Institutes for Research® (AIR®) and have experience with coaching school leaders reviewed the survey at two points during the design process to ensure that principals would interpret items as intended.[9]

AIR also created a second, shorter survey focused on the sustainability of CI, which was intended for school leaders from Cohort 1 and Cohort 2 networks that no longer received direct foundation support in the 2024–25 school year.

NSI School Leader Survey Data Collection

The study team administered the survey to school leaders annually. For the first administration (School Year 2020–21), the study team administered the survey in May. In subsequent school years, the study team worked with intermediary staff to identify survey windows in January through April. The intended sample was leaders from all schools that participated in the initiative, including the preidentified comparison schools.[10] The study team worked with staff from intermediaries to identify appropriate survey administration windows for each NSI, as well as to identify the respondent for each school. Intermediary staff helped promote the survey before launch and assisted with follow-up to improve response rates.

[9] These reviewers were not otherwise associated with the NSI evaluation; neither were they involved in AIR's work implementing NSI-funded networks in Florida and California.

[10] Evaluators at Mathematica Policy Research constructed this pool of comparison schools to estimate the impact of participation in NSI on student achievement outcomes. The comparison schools included all schools randomly assigned to serve as comparisons for Cohort 3 schools, as well as schools identified through Mahalanobis distance matching as comparisons for schools that had eighth-grade on-track or ninth-grade on-track outcomes.

Surveys were programmed and administered using Voxco survey software. Voxco allowed for complex skip logic (i.e., respondents saw only relevant questions based on their previous answers) as well as multiple opportunities for piped text (i.e., we piped in the name of the school's specific network as well as the grade on which the school's network focused). The emails that school leaders received about the survey also incorporated this piped-in text.

Each school leader received an email invitation to participate in the survey that referenced previous communication sent by intermediary staff. We then sent weekly or twice-weekly automated email reminders to encourage staff to participate. When email messages were unsuccessful, study team staff called schools to remind leaders about the survey.

Each respondent who completed the survey received a $50 gift card in appreciation of their participation.[11]

Response Rate

For response rates by NSI, see Exhibits C7, C8, C9, C10, and C11 for the primary school leader survey and C13 for the sustainability survey.

Intermediary Interviews

Intermediary Interview Design

The study team designed the intermediary interviews to collect information about each NSI, including their approach to supporting CI team activities. Intermediary interviews addressed the design, launch, and ongoing implementation of each NSI's CI school or district teams. Interview topics included the following:

- Intermediaries' approaches to developing a theory of practice improvement
- Intermediaries' approaches to identifying and selecting change ideas
- Data and measurement used in CI processes
- CI team membership
- Knowledge management

Using both a preinterview form for intermediary staff to complete and a structured protocol, the study team held interviews with key intermediary staff during the spring, focusing primarily on CI activities from the current school year.

[11] This incentive was not provided in districts where direct payment to research participants was not allowed. In addition, the amount offered was sometimes less if district research offices placed a limit on incentive payments at less than $50.

Intermediary Interview Data Collection

The interviews were conducted with intermediary staff who were familiar with districts' or schools' CI work and regularly interacted with CI teams, such as the CI lead, CI coach, and CI data specialist or lead. The protocol and preinterview form incorporated relevant information from existing data (e.g., intermediaries' applications, annual reports, information provided to the foundation and collected by other research or support partner teams) and prior interview data, as appropriate. Preinterview forms asked closed-ended (categorical or numerical) questions about CI teams' activities; the interviews asked open-ended questions about key CI team processes and decisions.

The study team reviewed the form data before conducting interviews and customized each interview protocol to reflect intermediaries' responses. For example, a study team member would pose different probing interview questions to an intermediary who had indicated on their form that conducting a root cause analysis was not part of the network's CI process than to an intermediary who had indicated that conducting a root cause analysis was part of their CI process. Each interview was conducted via Zoom, lasted from 60 to 120 minutes, and was recorded after obtaining interviewees' oral consent. The interviews included between two and eight intermediary staff, participating as a group.

In 2020–21, the team conducted interviews with 24 NSI between May and June 2021; we did not interview Partner MGIT-2 that year because the NSI had not yet been launched. In 2021–22, the study team conducted interviews with 13 NSI between June and July 2023. In 2022–23, the team conducted interviews with 15 NSI between April and May 2023. In 2023–24, the team conducted interviews with 10 NSI between April and May 2024. In 2024–25, the team conducted interviews with nine NSI in April 2025. See Exhibit C5 for the sample by year.

In general, the study team conducted one interview per NSI. Thus, intermediaries with two NSI generally completed two interviews. For example, the Network for College Success received two grants and operates two NSI, so it participated in two interviews. One exception occurred during the 2020–21 school year interviews: High Tech High, which operates two NSI, participated in one interview. The study team asked the CI leads the same questions during the single interview, and the responses were identified by NSI.

Each interview was transcribed and stored on the study team's secure SharePoint site (along with a running record of all documents for each NSI available to the evaluation team).

Response Rate

The team administered preinterview forms to intermediaries and conducted interviews with intermediaries during the 2020–21, 2021–22, 2022–23, 2023–24, and 2024–25 school years. The response rate for the preinterview forms and interviews was 100% in all 5 years. We did not collect landscape data from AIR Long Beach-3 in Year 5 because the network did not operate as it had in previous years.

Climate Surveys

Climate Survey Data Collection

To obtain data about perceptions of school climate, we drew on climate surveys routinely administered to students, teachers, and other stakeholders in districts in which NSI schools were located. Our evaluation partners at Mathematica Policy Research collected respondent- and school-level data about school climate from participating districts.

We limited the sample of districts that administered climate surveys to those that met the following conditions:

- The survey must have been administered while the network was active.
- The survey must have been administered to at least three schools in a district.
- To make it possible to pool schools using different surveys in a common analysis, the surveys needed to contain similar items and constructs. We used the University of Chicago Consortium's 5 Essentials for School Improvement framework (which is used in Chicago Public Schools) as the base to which we compared other districts' surveys. See Appendix E for more detail on the items and constructs in each survey.

Exhibit D3 shows the districts and NSI from which we collected data.

Exhibit D3. Districts From Which Climate Data Were Obtained

District	NSI
Chicago Public Schools	Network for College Success-1, Network for College Success-3, Teach Plus-3
Fresno Unified School District	CORE-1
Long Beach Unified School District	AIR-Long Beach, CORE-3
New York City Department of Education	New Visions-1, New Visions-3, Teaching Matters-3, Bank Street-3
Philadelphia Schools	Partners MGS1-2, Partners OTSN-2

Note. Data were also collected for Baltimore. However, data were not available at the item level to facilitate a crosswalk to the 5 Essentials for School Improvement framework.

Response Rate

For response rates by NSI, see Exhibit C12.

CI Team Rosters

Roster Data Collection

To obtain information about the stability of schools' CI teams, we use twice-annual collections of rosters conducted by our evaluation partners at the RAND Corporation. RAND collected information from each NSI twice a year (once in September and once in January) to document the schools involved in the initiative and the staff involved at the intermediary, district, and school levels. This process started in September 2021 and continued through January 2025.

Once we had received the rosters from RAND, we standardized individual team members' names and email addresses so that we could link them across time points. We then created indicators for each person to determine whether they were present in both the spring and fall of a specific school year, as well as from spring to spring of subsequent school years.

Response Rate

These data were complete for all schools.

Case Study Interviews

Case Study Interview Design

Case study interviews were conducted to gain an in-depth understanding of how CI processes occur within schools. Protocols for the NSI school case study interviews in 2020–21 were designed to collect data in 2020–21 in five areas:

1. Context, including teacher collaboration and family involvement
2. The CI team, including team selection, available resources, and the collaborative process
3. The implementation of school inquiry cycles
4. The adoption of tested change ideas, their equity focus, and diffusion beyond the CI team
5. The effect of the COVID-19 pandemic on CI processes

After reflecting on the 2020–21 results, the team refined the protocols to focus more specifically on inquiry cycles. The updated protocols for the 2021–22 through 2024–25 interviews aimed to collect data in three areas:

1. Decision-making around change ideas to test
2. Planning and implementation of change ideas in practice
3. Measurement and actions taken following implementation

Up to four interviews were conducted per school with staff familiar with CI work.

Case Study Interview Sample Selection

The study team conducted case studies in the same NSI included in RAND's evaluation of network implementation. The goal was to maximize learning by selecting strategically diverse NSI based on constructs theorized to affect CI implementation. The selection of the sample involved six steps:

Step 1: Identifying constructs. The study team identified the following six constructs theorized to affect the implementation of the CI process and on which the NSI were expected to differ in their approach, based on data from the intermediary interviews:

1. Commitment to CI
2. Diversity, equity, and inclusion integration
3. The intended role of school leaders
4. The prescriptiveness of intermediary guidance
5. School involvement in developing change ideas
6. The structure of the intermediary

Step 2: Maximizing learning about the constructs of interest. Two constructs per NSI were selected to ensure equal representation and varied approaches. For example, the study team assigned to NSI the construct "school involvement in developing change ideas" to represent high, medium, and low involvement levels.

Step 3: Selecting the number of schools per NSI to study. In 2020–21, 19 schools from seven NSI were selected based on NSI size and construct balance.

In 2021–22 and 2022–23, two schools from each of five NSI were selected.

In 2023–24 and 2024–25, two schools from each of four NSI were selected. See Step 6 below for further explanation.

Step 4: Identifying schools for the sample. In 2020–21, the study team scheduled brief phone calls with each intermediary to discuss variation across their schools in the two constructs of interest selected. For each construct, the intermediary was asked to think of schools that varied in their implementation of the construct. For example, if an NSI's construct was "intended role of school leader," the intermediary was asked to identify multiple schools in which the school leader was actively engaged in the CI process and multiple schools in which the leader was less involved.

Step 5: Selecting the schools in the 2020–21 sample. In 2020–21, the study team used the list of schools provided by the intermediary to purposely select a sample of schools that represented variation across four variables: urbanicity, school size, student demographics, and student outcomes.

Step 6: Identifying a subset of NSI in the sample. Of the seven NSI studied in 2020–21, five were selected to represent each cohort in 2021–22 and 2022–23. Two of the selected NSI were from Cohort 3 to enable the team to maximize the number of schools that could be studied from the point at which they had begun CI work. The intention was to maintain continuity by studying the same schools originally included in each of the five selected NSI, enabling the team to assess sustainability over time. In three cases, different schools were substituted at the request of the intermediary.

In 2023–24, Cohort 1 schools were excluded, because their grant had recently concluded. Instead, the study team reengaged these schools in the 2024–25 interviews to assess sustainability. One Cohort 3 intermediary was removed in 2024–25, leaving one intermediary in Cohort 1, two in Cohort 1B/2, and one in Cohort 3.

Case Study Interview Data Collection

Interviews were conducted via Zoom, primarily in spring, with some in summer or fall. Each school had one to four interviews, depending on staff availability and evolving study needs. All interviews were recorded and transcribed.

Response Rate

Exhibit D4. Response Rates for Case Study Interviews

Interview year	Number of schools interviewed	Number of schools in sample	Number of NSI in sample
2020–21	16	19	7
2021–22	10	10	5
2022–23	9	10	5
2023–24	7	8	4
2024–25	8	8	4

Note. NSI = Network for School Improvement.

Appendix E. Analysis Methods

The report addresses the following research questions (RQs) about the implementation of continuous improvement (CI):

- To what extent do schools implement disciplined inquiry cycles?
- To what extent do the activities undertaken reflect the core parameters and other evidence-based practices of CI?[12]
- What change ideas do schools select for testing and implementation as part of CI?
- To what extent do CI teams explicitly focus on improving outcomes for Black students, Latino students, or students experiencing poverty?
- Does the level of implementation of key CI features differ by cohort, late entry, outcome area, or grant year?
- What enabling conditions support schools' implementation of CI?
- To what extent is the level of CI implementation associated with school-level adoption and diffusion of effective interventions, diffusion of CI practices, and changes in school climate?
- To what extent do schools sustain CI after foundation funding ends, and what school-level characteristics are associated with sustainability?

The following sections describe the process to create measures used to address these questions and the analysis methods for each RQ.

Measure Development

We created measures based on the CI artifacts, school leader survey, school climate data, and the sustainability survey. Across these data sources, we constructed measures of disciplined inquiry cycles, implementation of core parameters, dimensions of equity, enabling conditions, CI diffusion, school climate, and CI sustainability. This section describes the process of creating each of these measures.

Measurement of Disciplined Inquiry Cycles

Measures of disciplined inquiry cycles were based on counts of cycle-level observations. We developed five school-level measures based on counts of cycle-level data: the number of cycles initiated, the number completed, the percentage completed, at least one cycle initiated, and the adoption of change ideas (Exhibit E1).

[12] The core parameters include: Root Cause Analysis, Aim Statement, Theory of Practice Improvement, Disciplined Inquiry Cycles, Teams, and Data Use/

Exhibit E1. Measures of Disciplined Inquiry Cycles

Measure	Definition
Number of cycles initiated	Number of cycles with evidence of at least one plan-do-study-act (PDSA) phase
Number of cycles completed	Number of cycles with evidence of all four PDSA phases
At least one cycle initiated	Binary indicator of whether a school had initiated at least one cycle
Percentage of cycles completed	The number of cycles completed as a percentage of the number of cycles initiated
Adopted a change idea	An indicator that determines whether a school reached the act phase, and if they did, whether they adopted a change idea in at least one cycle

To determine the number of cycles initiated and completed at the school level, we computed the number of cycles for each CI team in the school, and then we calculated the equally weighted average count of cycles for CI teams. Thus, the results should be interpreted as the number of cycles on average per school team. Most schools (94%) had only one CI team.

The measures of the number of cycles initiated and completed were restricted to schools that had at least one team that had initiated at least one cycle, based on Phase 2 artifact data. We truncated these school cycle averages to 10 for cycles initiated and completed, because in years 2022–23 and beyond, coders coded a maximum of 10 cycles per team.

We measured whether a school initiated at least one cycle based on whether a school provided Phase 2 artifact data. If it provided any Phase 2 artifacts, we considered that it initiated a cycle. If it did not provide Phase 2 artifacts but provided Phase 3 artifacts, we coded the school as not initiating at least one cycle. We treated schools as missing if they had not provided Phase 3 artifact data or our intermediary interviews indicated that cycles were conducted, but we lacked Phase 2 artifacts.

To determine the percentage of cycles that were complete, we divided the total number of cycles completed by the number of cycles initiated and multiplied the result by 100. For schools with more than one CI team, we computed the equally weighted average percentage of cycles completed among all CI teams in a school.

Finally, we created a measure indicating whether a school adopted at least one change idea during the act stage of an inquiry cycle.

Summary statistics for the cycle count measures are presented in Exhibit E2.

Exhibit E2. Summary Statistics Measures of Disciplined Inquiry Cycles

Measure	*N*	Mean	*SD*	Min	Max
Number of cycles initiated	725	3.37	2.50	1	10
Number of cycles completed	725	1.82	2.14	0	10
At least one cycle initiated	1,044	0.69	0.46	0	1
Percentage of cycles completed	1,222	68.90	46.30	0	1
Adopted a change idea	725	0.71	0.74	0	2

Note. Includes all years with artifacts available (see yellow and green shaded areas in Exhibit C1). *N* of NSI is 25.

Measurement of CI implementation

We created measures to assess the implementation of other five core parameters, listed in Exhibit E3.

Exhibit E3. Measures of CI Implementation of Core Parameters

Measure	Definition
Use of structured CI tools	Evidence that teams used structured tools and/or templates during inquiry cycles to implement, test, and evaluate changes
Use of data	Number of data use elements
Aim statement	Number of aim statement elements
Theory of practice improvement	Number of theory of practice improvement elements
Root cause analysis	Number of root cause analysis elements

To create these measures, we analyzed artifact data using the Rasch rating scale model (Andrich, 1978),[13] as implemented in WINSTEPS version 5.4.1.0 (Linacre & Wright, 2026).[14] A rating scale model is a psychometric approach that calculates item difficulties to estimate a person's ability based on the selected items, where "person" refers to the entities being measured—in this case, schools. For clarity, in the discussion that follows, we will use "school" rather than person to refer to Rasch person statistics.

Applying the Rasch model enabled us to examine the psychometric properties of scales constructed from binary and ordinal items derived from the artifacts. The Rasch rating scale model is useful because it allows us to test whether a set of items from the artifacts measures a single latent variable.

To construct each measure, we selected a group of items intended to measure each core parameter and developed a Rasch model based on the items. The items selected in each measure are listed in Exhibit E4. The use of a structured tools measure is based on Phase 2 artifacts. Each of the four items is

[13] Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*(4), 561–573.
[14] Linacre, J. M., & Wright, B. D. (2026). Winsteps

binary, indicating that structured protocols were used in a particular stage of a plan-do-study-act cycle. The measure is missing for schools that did not conduct cycles.

The other four measures are based on Phase 3 artifacts. An item was coded 1 if the defined element was observed in the Phase 3 artifacts for a school. It was coded 0 if Phase 3 artifacts were provided for a school, but the element was not observed, and missing if no Phase 3 artifacts had been provided for a school. Items are ordered by their difficulty, as calculated by the Rasch model (easier items are listed first).

Exhibit E4. Descriptions of CI Implementation Measures

Measure	Description	Item no.	Items (evidence of…)
Use of structured CI tools	Evidence that teams used structured tools and/or templates during inquiry cycles to implement, test, and evaluate changes	1	Structured protocols used in the plan stage
		2	Structured protocols used in the do stage
		3	Structured protocols used in the study stage
		4	Structured protocols used in the act stage
Use of data	Number of data use elements	1	Data use supported by intermediary
		2	Data used to develop plan
		3	Data to be collected align with plan
		4	Data collected or analyzed during study phase of inquiry cycle
Aim statement	Number of aim statement elements	1	Explicit aim statement
		2	“Relevant” criterion for SMART goal
		3	“Specific” criterion for SMART goal
		4	“Measurable” criterion for SMART goal
		5	“Attainable” criterion for SMART goal
		6	“Attention to equity in aim statement
		7	“Time-bound” criterion for SMART goal
Theory of practice improvement	Number of theory of practice improvement elements	1	Theory of practice improvement developed
		2	Driver diagram related to aim
		3	Change ideas related to drivers
		4	Based on school-specific data
		5	Attention to equity in theory of practice improvement
Root cause analysis	Number of root cause analysis elements	1	Root cause analysis (RCA) conducted
		2	At least one RCA activity completed (e.g., Five Whys, Influence Map, RCA surveys, Empathy Interviews, fishbone diagrams)
		3	Agenda, meeting notes, or structured inquiry cycles mentioning RCA
		4	Discussion of RCA (e.g., part of a training, RCA templates provided)
		5	RCA that demonstrates an understanding of factors that produce inequitable outcomes
		6	Attention to equity in the RCA

Note. SMART = specific, measurable, achievable, relevant, and time-bound.

To assess model fit, we examined the outfit, infit, reliability, and dimensionality of each measure. All fit statistics are based on suggestions from the WINSTEPS manual (Linacre, 2026),[15] which provides information on each fit statistic. A summary of fit statistics is in Exhibit E5.

The outfit mean-square statistic assesses whether the observed item responses are distributed around the responses predicted by the Rasch model as expected. The infit mean-square statistic is similar but gives less weight to outliers. Infit and outfit statistics that fall within a range of 0.8 and less than 1.2 generally indicate good fit.

Reliability statistics assess the internal consistency of a measure—how closely a set of items are related as a group. We report Cronbach's alpha, person reliability, and item reliability. Cronbach's alpha is a widely used measure of reliability not specific to the Rasch model, based on a scale that sums the items rather than apply the Rasch model. Person reliability indicates the reliability of the Rasch measure in distinguishing among the "persons" measured—in this case, schools. Item reliability indicates the reliability of the Rasch measure in distinguishing among the items in terms of difficulty. We considered adequate thresholds to be 0.7 for Cronbach's alpha, 0.5 for person reliability, and 0.9 for item reliability.

The Rasch model assumes the unidimensionality of a measure. Unidimensionality is an assumption that a set of items measures a single underlying construct. To assess the assumption of unidimensionality, we report the first eigenvalue for a principal components analysis of the residuals from the Rasch model. The eigenvalue assesses the variance explained by the largest potential secondary dimension. A Rasch model with a first eigenvalue below 2 indicates the model is likely unidimensional. In addition, we report the percent variance explained, which indicates how well the Rasch measure accounts for the variance in the data. We used a value of at least 50% variance explained to identify a model with good fit.

[15] Linacre, J. M. (2026). *Winsteps program manual* (Version 5.11.0). Winsteps.com. https://www.winsteps.com/a/Winsteps-Manual.pdf

Exhibit E5. Fit Indices and Reliability for Scaled CI Implementation Measures

Measure	Number of items	School measures				Item measures				Percent variance explained
		Cronbach's alpha	Infit	Outfit	Reliability	Infit	Outfit	Reliability	Eigenvalue	
Use of structured CI tools	4	0.78	1.01	0.95	0.37	0.99	0.95	0.99	1.67	38.80%
Use of data	4	0.58	0.91	1.05	0.43	0.95	1.74	1.00	1.36	58.10%
Aim statement	7	0.88	1.01	0.91	0.59	1.00	0.91	0.95	1.85	34.20%
Theory of practice improvement	5	0.76	0.91	0.95	0.67	0.96	1.12	1.00	1.83	58.20%
Root cause analysis	6	0.92	1.02	0.79	0.82	0.98	0.79	0.97	1.82	34.80%

After confirming model fit, we generated school-level Rasch scale scores, scaled to the number of items in the measure. Exhibit E6 shows summary statistics for each of the CI implementation measures.

Exhibit E6. Mean and *SD* of CI Implementation Measures

Measure	*N*	Mean	*SD*	Min	Max
Use of structured CI tools	725	2.66	1.40	0	4
Use of data	1,222	1.69	1.21	0	4
Aim statement	1,222	4.50	2.14	1	7
Theory of practice improvement	1,222	2.69	1.61	0	5
Root cause analysis	1,222	2.37	2.50	0	6

Measurement of Equity Dimensions

To measure the equity dimensions of CI implementation, we drew on the equity framework described in Box 11 of the main report. We identified artifact items that aligned with each of the four dimensions: access, achievement, identity, and agency. We used five items to construct a measure for each dimension, with one item for each of five core parameters (disciplined inquiry cycles, use of data, root cause analysis, aim statement, and theory of practice improvement). Each item was coded 1 if there was evidence in the artifacts that the relevant core parameter (e.g., disciplined inquiry cycles) occurred and the dimension of equity was present for that parameter. It was coded 0 if the core parameter did not occur or evidence for the dimension of equity was not present. The Items used to construct each measure are shown in Exhibit E7.

Exhibit E7. Measures of Dimensions of Equity

Measure	Definition
Access	Evidence that teams focused on equity by providing resources to students, including tangible and intangible assets and opportunities
Achievement	Evidence that teams focused on equity by focusing on improved student outcomes, particularly students within historically marginalized groups
Agency	Evidence that teams focused on equity by providing resources and interventions that provide students from nondominant backgrounds an increased opportunity to use their voices and express agency to challenge contemporary inequalities within and beyond school walls with the goal of engendering structural change
Identity	Evidence that teams focused on equity by providing resources and interventions that are attentive to a student or teacher's background.
Overall	The average of the four equity dimension measures

Each measure was calculated as the sum of the five items, creating a 0–5 scale that counted evidence of equity across all five core parameters for each dimension. We attempted to scale the measures by using a Rasch model, but the model fit was inadequate. This could be because not all schools had engaged in all five core parameters, or because some equity items had means near zero. We also created an overall equity measure, which was an average of all four equity dimension measures. Summary statistics for the equity dimension measures are presented in Exhibit E8.

Exhibit E8. Summary Statistics for Equity Dimension Measures

Measure	*N*	Mean	*SD*	Min	Max
Access	1,222	0.91	1.13	0	4
Achievement	1,222	1.27	1.17	0	5
Agency	1,222	0.57	0.82	0	4
Identity	1,222	0.50	0.78	0	4
Overall equity	1,222	0.81	0.71	0	3.50

Measurement of Enabling Conditions

Measures of enabling conditions were based on items from the NSI school leader survey. To construct these measures, we used the same Rasch modeling approach used to construct the CI implementation measures described above. The items used for each measure are shown in Exhibit E9. In one case, for the school leader involvement scale, items had different response scale options. Therefore, rather than use the rating scale model, we used a partial credit model, which allows each item to have its own response structure. Two measures, planning time and CI team meeting frequency, were single item measures, and therefore did not use a Rasch scale.

Exhibit E9. Enabling Conditions Measures

Measure	Description	Item no.	Survey item	Response options
School leader involvement	Level of school leader involvement in continuous improvement (CI)	1	How involved are you, as principal, CURRENTLY with the [Network Name]?	Not at all involved (1) Minimally involved (e.g., I am aware of what is happening but do not actively participate or make major decisions) (2) Moderately involved (e.g., I occasionally attend our continuous improvement team meetings, am sometimes involved in key decisions, or attend some of the network events) (3) Very involved (e.g., I attend most continuous improvement team meetings, but I do not lead the team, and I attend most network events) (4) Very actively involved (e.g., I actively lead or facilitate the team leading our continuous improvement work) (5)
		2	This school year, how are you involved in these educators' collaborative meetings?	Neither (0) I attend some of the meetings (1) I attend all of the meetings (2)
School supports	Changes made to the school that included hiring new staff, changing staff roles, changing the curriculum, providing coaching, changing schedules, providing a stipend, creating a common planning period, reducing teacher loads, and providing time during afterschool meetings	1	Has your school made any of the following changes to facilitate participation in the [Network Name]? Options are: Yes/No	Hired new staff
		2		Changed staff roles to accommodate new responsibilities
		3		Reduced teacher loads to allow for more planning time
		4		Changed the school schedule
		5		Changed the math curriculum
		6		Changed the ELA curriculum
		7		Provided targeted professional development
		8		Provided targeted coaching
		9		Provided a stipend or compensation to staff participating
		10		Provided time during afterschool staff meetings
		11		Created a common planning period for staff on the team
CI team meeting frequency	Teachers meet at least every 2 weeks as part of a collaborative group.	1	This school year, how do you monitor these educators'	At least two times per week (1) Weekly (1) Every two weeks (1)

Measure	Description	Item no.	Survey item	Response options
			collaborative meetings? (Select all that apply.)	Monthly (0) Quarterly (0) It varies (0)
Planning time	Teachers have at least 5 hours of planning time in their contract.	1	How much individual planning time do [focal grade] teachers at your school have per week?	
Network for School Improvement (NSI) support for knowledge management	Level of supports related to CI knowledge management provided by the intermediary	1	[INTERMEDIARY] provided processes and tools that my school's [team name] team uses to … Options are: Strongly Disagree (1), Disagree (2), Agree (3), Strongly Agree (4)	document their continuous improvement activities.
		2		review their continuous improvement activities.
		3		share what they have learned through their continuous improvement activities with other staff in our school.
		4		share what they have learned through their continuous improvement activities with other schools.
		5		review what other schools' CI teams have learned from their continuous improvement activities.
NSI supports to schools	Level of support the network and intermediary provide to schools	1	At our school … Options are: Strongly Disagree (1), Disagree (2), Agree (3), Strongly Agree (4)	[Intermediary name] has supported our school's use of continuous improvement strategies.
		2		[Intermediary name] provides valuable support to our school and students.
		3		[Intermediary name] supports my school in accessing the relevant data and effectively using data for continuous improvement cycles.
		4		The [Network Name] gives my staff the opportunity to participate in training or consultation regarding evidence-based interventions and/or best practices.
		5		The [Network Name] has provided me opportunities to connect with and learn from other schools/principals.
		6		The [Network Name] has provided my staff opportunities to connect with and learn from educators at other schools.

We use the fit statistic values described above for the CI implementation measures to assess the fit of the enabling condition measures. Exhibit E10 summarizes the fit statistics for each measure.

Exhibit E10. Fit Indices and Reliability for Scaled Enabling Condition Measures

Measure	Number of items	School measures				Item measures				Percent variance explained
		Cronbach's alpha	Infit	Outfit	Reliability	Infit	Outfit	Reliability	Eigenvalue	
School leader involvement	2	0.83	0.88	0.97	0.60	0.97	0.98	0.98	N/A	64.50%
School supports	11	0.72	1.00	1.01	0.52	0.99	1.02	0.99	1.67	32.60%
NSI support for knowledge management	5	0.92	0.72	0.78	0.75	0.96	0.84	0.99	2.64	61.30%
NSI supports for schools	6	0.92	0.82	0.85	0.79	0.98	0.86	0.96	1.97	55.00%

We generated school-level Rasch scale scores scaled to the original scale of the items. The specific approach differed for the four measures. Two measures were based on Likert scale items (NSI support for knowledge management systems and NSI supports to schools). For these, we scaled the value from 1–4 to represent the Likert options (1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree). School supports used binary items, so we scaled the measure to the number of items, 11. School leader involvement used a partial credit model, for which we scaled the measure to the total number of item options (1–7).

Two enabling condition measures used the school roster data to measure team stability. We calculated cross-year and within-year team stability to measure the extent to which CI teams were consistent within and across years. Exhibit E11 defines each of these team stability measures. Exhibit E12 presents summary statistics for the enabling conditions measures.

Exhibit E11. Team Stability Measures

Measure	Definition
Cross-year team stability	Percentage of members on a CI team that stayed the same over the course of the school year and from year to year
Within-year team stability	Percentage of members on a CI team that stayed the same from fall to spring within the same year

Exhibit E12. Summary Statistics for Enabling Conditions

Measure	*N*	Mean	*SD*	Min	Max
School leader involvement	941	3.06	1.44	0	6
School supports	830	3.70	2.19	0	11
Continuous improvement team meeting frequency	755	0.71	0.45	0	1
Planning time	773	0.48	0.50	0	1
NSI support for knowledge management	833	2.98	0.58	1	4
NSI supports for schools	834	3.12	0.59	1	4
Cross-year team stability	563	0.65	0.29	0	1
Within-year team stability	841	0.91	0.19	0	1

Note. NSI = Networks for System Improvement.

Measurement of School Climate

We created climate measures for two purposes: first, to serve as enabling conditions (for RQ6); and second, to assess whether the implementation of CI led to improvements in school climate (for RQ7). The climate measures were based on climate survey data obtained from a sample of study districts that had already administered school climate surveys (see Appendix D).

To facilitate analysis across districts that used different school climate survey instruments, the team created a crosswalk of items from each of the district teacher surveys to identify items similar to those found in the 5 Essentials for School Improvement used in Chicago. (Although the surveys included similar items, the organization of the items into constructs and subconstructs differed across surveys. Thus, we focused on identifying similar items, and we used the items we identified to create constructs parallel to those in the 5 Essentials for School Improvement). For the purposes of the analyses in this report, we focused on two essentials: *collaborative teachers* and *effective leaders*. For both these measures, we relied exclusively on school-level data summarizing teachers' responses to climate surveys. Exhibit E13 displays the subconstructs contained within each essential and the number of relevant survey items from each district measuring each subconstruct.

Exhibit E13. Number of Survey Items for Each Climate Subconstruct, by District

Essential	Subconstruct	Chicago	Philadelphia	New York City	Long Beach and Fresno
Collaborative teachers	Collaborative practices	5	4	1	0
	Collective responsibilities	6	2	2	1
	Quality professional development	5	4	4	0
	School commitment	4	0	2	0
	Teacher–teacher trust	5	1	5	3
Effective leaders	Instructional leadership	6	4	6	0
	Program coherence	5	2	2	0
	Teacher influence	5	4	3	0
	Teacher–principal trust	8	2	8	0

We had item-level data available for New York City, Philadelphia, and the CORE Districts (Fresno and Long Beach). These data sets provided the number and percentage of teachers who responded to each option on a Likert scale (e.g., *strongly agree* or *agree*). Chicago Public Schools provided school-level scale scores for each subconstruct.

In the districts where we had item-level information, we first created a school-level weighted average response for each item. For example, if a response scale consisted of four options (*strongly disagree*, *disagree*, *agree*, and *strongly agree*), we multiplied the number of respondents who selected *strongly agree* by four, we multiplied the number of respondents who selected *agree* by three, and so forth. We added these products together and then divided by the total number of respondents to create an average score. So, an average score of 3.5 would indicate that the teachers in that school provided an average response between *strongly agree* and *agree*. We then averaged each of these item scores to create scores for each subconstruct (e.g., teacher–teacher trust or program coherence). We then averaged each of the subconstructs to generate the overall measure score (e.g., collaborative teachers or effective leaders). We equally weighed subconstructs even if they had different numbers of items. In Chicago, where we did not have item-level data, we averaged each of the subconstruct scale scores within each measure to create overall scores of collaborative teachers and effective leaders. [16]

We then standardized each subconstruct and measure score within district and within year. This gave us a score for each school that represented the school's position within the districtwide distribution each year. Doing so facilitated cross-year and cross-district comparisons of schools' measures of climate by allowing us to compare schools based on their distance from their district average rather than on their raw scores.

[16] Information about the reliability of the scales is available upon request.

To calculate the change in school climate used in RQ7, we subtracted the value in Grant Year 1 from Grant Year 3 for each school.

Exhibit E14 displays the summary statistics for the school climate measures.

Exhibit E14. Summary Statistics for Measures of Climate

Measure	*N*	Mean	*SD*	Min	Max
Collaborative teachers	305	0.04	0.86	-2.23	2.29
Effective leaders	289	0.06	0.96	-3.39	2.40
Change in collaborative teachers, Grant Years 1–3	105	0.01	0.72	-1.96	2.21
Change in effective leaders, Grant Years 1–33	98	-0.01	0.94	-2.50	2.41

Measurement of CI Diffusion

For measures of CI diffusion (used in RQ7), we used single-item measures derived from the NSI school leader survey. These measures are described in Exhibit E15. Summary statistics are presented in Exhibit E16.

Exhibit E15. Description of CI Diffusion and Sustainability Measures

Construct	Measure	Item/description
Diffusion of CI practices	Other staff adopted CI practices	Other staff or teams in my school have adopted the CI practices enacted by our school's CI team. (Yes = 1; No = 0)
School-level adoption and diffusion of effective interventions	Other staff adopted interventions	Other staff or teams in my school have adopted the strategies or interventions tested by our school's CI team. (Yes = 1; No = 0)

Exhibit E16. Summary Statistics for CI Diffusion Measures

Measure	*N*	Mean	*SD*	Min	Max
Diffusion of CI practices	239	0.75	0.43	0	1
Other staff adopted interventions	738	0.71	0.45	0	1

Measurement of Sustainability

We used single items from the school leader sustainability survey to report on the sustainability of the NSI initiative. Exhibit E17 lists the items used, and Exhibit 18 lists the summary statistics. Summary statistics in Exhibit E18 are not weighted by NSI and therefore may differ from averages in the main report.

Exhibit E17. Measures of Sustainability

Measure	Item/description	Response options
School continues to have a CI team	As a part of your school's involvement in [network], your school had a [team]. Is this team still functioning in your school?	**This team still meets and carries out the same responsibilities as before.** **This team still meets, but they have different responsibilities.** This team has been absorbed into another team/committee/collaborative group in our school. This team no longer meets, but some of what they were responsible for have been integrated into one or more other groups. The team no longer meets, and its work has not been integrated into other groups.
School continues to use a CI approach	As a part of the [network], the [intermediary] encouraged your school to use a specific approach to continuous improvement—for example, plan-do-study-act cycles or goal setting. Does your school continue to use this approach?	**Yes, our school is continuing to use this approach as the [intermediary] introduced it to us.** **Yes, our school is continuing to use the approach, but we modified it to make it easier for our school to use.** No, our school implemented the approach but is no longer using it. No, our school never implemented the approach the [intermediary] recommended. I'm not sure about the approach to CI the [intermediary] recommended.
Helped to continue the work of the CI team: Key staff in place	Have any of the following helped your school continue to implement the work of the [team]? Yes = 1; No = 0	key staff in place
Helped to continue the work of the CI team: Staff buy-in		Staff buy-in
Helped to continue the work of the CI team: Time and space for staff to collaborate		Time and space for staff to collaborate
Helped to continue the work of the CI team: Complementary school- or district-based initiatives		Complementary school- or district-based initiatives
Helped to continue the work of the CI team: Administrator stability		Administrator stability
Helped to continue the work of the CI team: Adequate funding		Adequate funding
Helped to continue the work of the CI team: Systems to document CI processes		Systems to document continuous improvement processes
Helped to continue the work of the CI team: Templates to guide the work provided by [intermediary]		Templates to guide the work provided by [Intermediary]

Exhibit E18. Summary Statistics for Sustainability Items

Item	*N*	Mean	*SD*	Min	Max
School continues to have a CI team	93	0.48	0.50	0	1
School continues to use a CI approach	87	0.38	0.49	0	1
School kept *key staff in place* to facilitate continued CI work	63	0.84	0.37	0	1
School *administrator stability* facilitated continued CI work	63	0.62	0.49	0	1
***Staff buy-in* facilitated continued CI work**	63	0.57	0.49	0	1
***Systems to document CI processes* facilitated continued CI work**	63	0.51	0.50	0	1
***Providing targeted professional development* helped sustain CI work**	63	0.48	0.50	0	1
***Changing staff roles to accommodate new responsibilities* helped sustain CI work**	63	0.39	0.49	0	1
***Providing time for staff to meet during the school day* helped sustain CI work**	63	0.37	0.49	0	1
***Providing time for staff to meet after the school day* helped sustain CI work**	63	0.33	0.48	0	1

Statistical Approach

The sections that follow describe the methods used for each RQ.

Research Questions 1–4

To address RQs 1–4, we focus on Grant Year 5 for each NSI: 2022–23 for Cohort 1; 2023–24 for Cohort 1B/2; and 2024–25 for Cohort 3.

We examine, by NSI, the five measures of disciplined inquiry cycles, five measures of CI implementation, and five measures of equity dimensions (15 in all) described above. To compute the average of each measure for each NSI, we computed the equally weighted mean for NSI schools with data on the measure in Grant Year 5. The overall mean was computed as the equally weighted mean for the NSI included in the analysis.

Research Question 5

RQ5 focuses on the relationship between network and school characteristics and the implementation of CI, as tapped by the 15 measures used in RQs 1–4. RQ5 also focuses on change in implementation over time. All analyses except those for change over time include all available years of data for all schools. The analyses for change over time are limited to NSI with data for at least one school in both Grant Years 3 and 5, because we lacked data for Cohort 1 in Grant Year 1, which took place in 2018–19, before data collection began. See Appendix Table C1 for data availability by cohort.

Relationship between network and school characteristics and CI implementation. To assess the extent to which the implementation of CI varied across NSI and schools within NSI, and to examine the extent to which implementation was associated with cohort, outcome area, and late entry, we estimated multilevel models, with school time points nested within schools, and schools nested within NSI. The model includes a random intercept for NSI to account for the correlation among schools in the same network, as well as a random intercept for the school to account for the correlations among observations across multiple years in the same school. The models can be expressed as follows:

$$Y_{ijk} = \beta_0 + \beta_1(LateEntry_{jk}) + \sum_s \beta_s OutcomeArea_k + \sum_c \beta_c Cohort_k + \sum_g \beta_g GrantYear_{ijk} + u_{0k} + v_{0jk} + e_{ijk}$$

In this model:

- Y_{ijk} represents one of the 14 CI implementation outcome measures, where i is an index for school time points, j is an index for school, and k is an index for NSI.
- $LateEntry_{jk}$ is a binary school-level characteristic for school j in NSI k.
- $\sum_s \beta_s OutcomeArea_k$ are binary NSI-level variables representing outcome area fixed effects, where one outcome area is omitted as the reference category. s represents the number of nonomitted student outcome areas. Outcome area coefficients (eighth grade on track, ninth grade on track, college-ready on track) are measured in reference to well-matched postsecondary, the omitted category.
- $\sum_c \beta_c Cohort_k$ are binary NSI-level variables representing cohort fixed effects, with one cohort omitted as the reference category. c represents the number of non-omitted cohorts. Cohort 2 and 3 coefficients are measured in reference to Cohort 1, the omitted category.
- $\sum_g \beta_g GrantYear_{ijk}$ are binary variables representing grant year fixed effects, with one grant year omitted as the reference category. g represents the number of nonomitted cohorts. There are three grant years in the model: Grant Year 1/2, Grant Year 3, and Grant Year 5. Grant Year 1/2 and Grant Year 5 are measured in reference to Grant Year 3, the omitted category.
- u_{0k} represents the random intercept for NSI k, which captures the unobserved differences between NSI. We assume $u_{0k} \sim N(0, \sigma_u^2)$.
- v_{0jk} represents the random intercept for school j within NSI k, which captures the unobserved differences between schools after accounting for variation between NSI. We assume $v_{0jk} \sim N(0, \sigma_v^2)$.
- e_{ijk} is the residual or error term capturing variation in the outcome that is not explained by the model covariates or the random effects.

To summarize the results, we reported coefficients and *p*-values for each variable of interest (Exhibit F1). We calculated a joint Wald test for both student outcome area fixed effects and cohort fixed effects, separately. We did not report coefficients for the grant year fixed effects, because we focused on the effect of grant year in the analysis of change over time.

Change over time. To report on how CI implementation in schools changed over time, we analyzed the relationship between grant year and the CI implementation measures. We employed a multilevel model similar to the one used to analyze the relationships between network and school characteristics with a few modifications. To isolate the effect of grant year for different outcome areas, we included only grant year and outcome domain as covariates. In addition, we included a random slope with an unstructured covariance matrix, meaning that the random intercepts and slopes may correlate without restriction. The models can be expressed as:

$$Y_{ijk} = \beta_0 + \beta_1(GrantYear_{ijk}) + \sum_s \beta_s OutcomeArea_k + u_{0k} + \sum_g u_{gk} GrantYear_{ijk} + v_{0jk} + e_{ijk}$$

In this model:

- Y_{ijk} represents one of the 14 CI implementation outcome measures, where i is an index for school time points, j is an index for school, and k is an index for NSI.
- $\sum_s \beta_s OutcomeArea_k$ are binary NSI-level variables representing outcome area fixed effects, where one outcome area is omitted as the reference category. s represents the number of non-omitted student outcome areas. The outcome area coefficients (eighth grade on track, ninth grade on track, college-ready on track) are measured in reference to well-matched postsecondary, the omitted category.
- $\beta_1(GrantYear_{ijk})$ is the binary variable representing grant year. Grant Year 3 is omitted as the reference variable, with one indicator for Grant Year 5.
- u_{0k} represents the random intercept for NSI k, which captures the unobserved differences between NSI. We assume $u_{0k} \sim N(0, \Sigma_u)$, indicating that the NSI-level random intercept may correlate freely with the NSI-level slope.
- $\sum_g u_{gk} GrantYear_{ijk}$ represents the random slope for grant year. We assume $u_{0k} \sim N(0, \Sigma_u)$, indicating that the NSI-level random intercept may correlate freely with the NSI-level intercept.
- v_{0jk} represents the random intercept for school j within NSI k, which captures the unobserved differences between schools after accounting for variation between NSI. We assume $v_{0jk} \sim N(0, \sigma_v^2)$.
- e_{ijk} is the residual or error term capturing variation in the outcome that is not explained by the model covariates or the random effects.

To summarize the results, we reported coefficients and *p*-values for the Grant Year 5 indicator (Exhibit F2). We also report the random effects parameters. To calculate NSI-level predictions for the exhibits, we calculated the predicted value of each NSI's average school for Grant Year 3 and Grant Year 5 separately. The overall mean is the average of the predicted values of each NSI weighted equally.

Research Question 6

RQ6 includes descriptive analyses focused on the average levels of enabling conditions for each NSI, as well as analysis of the relationship between CI implementation levels and enabling conditions.

Descriptive analyses. For the descriptive analysis, we focused on the 10 measures of enabling conditions described in the measurement section above. We report average measures by NSI for each enabling condition, organized by outcome area. The analyses focused on Grant Year 5, except for the school climate measures, which focused on Grant Year 3 due to lack of availability of data for all schools in Grant Year 5. To compute averages of measures by NSI, we computed the equally weighted mean for the NSI schools included in the analysis. The overall mean was computed as the equally weighted mean for all NSI included in the analysis.

Relationship between CI implementation and enabling conditions. We employed the sample multilevel model as the one used in the previous section detailing the relationships between network and school characteristics and CI implementation (RQ5), with the addition of each of the 10 enabling condition measures to a series of models as an independent variable. The sample was limited to schools with measures enabling conditions and CI implementation for each model.

Because of the lack of variation in outcome area in the sample of schools with school climate data, we removed the fixed effects for outcome area for the models with school climate. We have 15 measures of CI implementation and 10 measures of enabling conditions, for a combined total of 150 models. We report on the main enabling condition coefficients and *p*-values in Tables F3, F4, and F5.

Research Question 7

RQ7 includes descriptive analyses focused on CI diffusion and change in school climate, as well as their relationship with CI implementation.

Descriptive analyses. For the descriptive analysis, we report average measures by NSI for the two measures of CI diffusion and the two measures of change in school climate described in the measurement section. To compute averages of measures by NSI, we computed the equally weighted mean for NSI schools included in the analysis. The overall mean was computed as the equally weighted mean for all NSI included in the analysis.

Relationship between CI implementation, diffusion, and change in climate. For this, we employed the same multilevel model as the one used in the "Relationships between network and school characteristics and CI implementation" section, with some changes. First, the 15 measures of CI

implementation were included as independent variables instead of as dependent variables. Second, the two measures of CI diffusion and the two measures of change in school climate were treated as the dependent variables. For the models with change in school climate, we do not have multiple time points, so we removed the random effect (v_{0jk}) for schools.

For the analysis of relationships between CI implementation and CI diffusion and change in school climate, the sample was limited to schools with measures of both CI implementation and the relevant CI diffusion or change in school climate measure. For the change in school climate, we focused on change between Grant Year 3 and Grant Year 1, because of the limited school climate data available in Grant Year 5.

We report on the main CI implementation coefficients and *p*-values in Appendix Table F6.

Research Question 8

RQ8 includes descriptive analyses focused on CI sustainability, as well as the relationship between enabling conditions, CI implementation, and sustainability.

Descriptive analysis. For the descriptive analysis, we summarized the two binary measures of sustainability—(1) the sustainability of the team and (2) the sustainability of practices—at the network level using model-based estimates from an empty random-intercepts mixed model. For each measure, we obtained predicted values that represent empirical Bayes network-level averages. Overall means were calculated as the equally weighted average of these network-level predicted values.

We also report on the percentage of respondents that reported various facilitators of sustainability and steps to sustain implementation. These percentages were computed as the unweighted mean of school-level responses.

Relationship between CI implementation measures, enabling conditions, and sustainability measures. We employed the same multilevel model as the one used in the "Relationships between network and school characteristics and CI implementation" section, with some changes. First, we used the 10 measures of enabling conditions and 15 measures of CI implementation as independent variables. Second, the two binary measures of sustainability were treated as dependent variables.

The sample for the analysis was limited to schools that participated in the sustainability survey following the end of the foundation grant and had at least some of the 10 enabling condition measures or 15 measures of CI implementation in Grant Year 5. Given the relatively small sample, we estimated an ordinary least squares regression with cluster robust standard errors rather than a multilevel model.

We report on the main enabling condition and CI implementation coefficients and *p*-values in Appendix Table F7.

Qualitative Analysis Methods

Intermediary Interview Data Analysis

The data from interviews were analyzed in two ways. First, intermediaries' responses to the preinterview form were coded and tabulated on the basis of each question's closed-ended response options. For example, the preinterview form asked, "Did your intermediary or participating CI teams test change ideas during School Year 20XX–XX?" and NSI chose between two responses, 'Yes' or 'No.' Change ideas were not tested in School Year 20XX–XX and will be in a future year." Second, interview transcripts were coded to capture information related to broad constructs (e.g., approach to engaging in CI). To analyze broad constructs, the study team employed an emergent coding approach to identify themes, patterns, and categories.

Case Study Interview Data Analysis

The study team analyzed all interview data and created diagrams to represent each school's CI process. For each interview, the team identified which components of CI were present—conducting a root cause analysis, setting an aim, selecting a change idea, preparing to implement the change idea, implementing activities, measuring results, and making decisions about next steps—and documented the order and way in which they were implemented. Each component was described in a labeled box, with arrows indicating the sequence of implementation. These visual representations provided a holistic view of each school's CI journey, highlighting variations in component use and order. For example, if a school selected a change idea without first conducting a root cause analysis, setting an aim, or measuring results, the diagram would reflect which components had been implemented. Comparing these visual representations across schools enabled the team to comprehensively describe variations in CI implementation.

To analyze data across all 5 years of the study, the team developed a detailed coding spreadsheet capturing key components: team structure, inquiry cycles, sustainability, postgrant activities, and other important areas. The team systematically reviewed each transcript from 2020–21 to 2024–25, documenting relevant information in the corresponding cells of the spreadsheet. Once all transcripts had been coded, a final column was used to synthesize information in each row, distilling the essence of each school's CI journey and how it had evolved over time across the identified components. The insights captured in the analysis column informed a comprehensive memo that profiled each school's unique approach to CI, highlighting patterns, shifts, and areas of emphasis throughout the 5-year period.

Appendix F. Results for Analysis of Variation and Relationship Between Enabling Conditions and the Implementation of Continuous Improvement

Exhibit F1. Analysis of the Relationship Between CI Implementation Measures and Cohort, Outcome Area, and Late Entry

Implementation measure	*n*	Covariates						Significance tests			ICC NSI
		Cohort		Outcome area				Wald test			
		Cohort 2	Cohort 3	8GOT	9GOT	CROT	Late entry	Cohort	Outcome area	Late entry *p*-value	
At least one cycle initiated	1044	-0.071	0.086	-0.159	-0.086	0.063	-0.114**	0.465	0.546	0.001	0.263
Number of cycles initiated	725	-0.725	-0.709	0.442	-1.146	-1.066	0.262	0.646	0.104	0.221	0.272
Number of cycles completed	725	-0.815	-0.446	1.336	0.129	1.160	-0.210	0.628	0.232	0.217	0.373
Percentage of cycles completed	725	-11.289	4.643	18.834	13.818	52.745	-10.72**	0.548	0.108	0.001	0.404
Change idea	725	-0.538	-0.083	0.229	0.171	1.348	-0.129	0.586	0.287	0.239	0.430
Use of structured CI tools	725	-0.246	-0.186	0.468	0.252	0.386	0.036	0.623	0.336	0.574	0.269
Root cause analysis	1222	0.227	0.298	-0.544	-1.109	-0.922	-1.108**	0.846	0.263	0.000	0.337
Aim statement	1222	-1.287	0.779	-1.057	-2.331	-0.413	-0.220	0.021*	0.093	0.132	0.375
Theory of practice improvement	1222	-0.334	0.965	-1.294	-1.390	-1.082	0.320**	0.011*	0.105	0.000	0.368
Use of data	1222	-0.164	0.552	-0.421	-0.288	0.349	-0.157*	0.088	0.368	0.042	0.306
Equity—access	928	0.435	1.153	-1.115	-0.910	-1.382	-0.025	0.402	0.008**	0.714	0.321
Equity—achievement	928	1.223	0.475	-0.027	0.182	0.140	-0.131	0.002**	0.901	0.065	0.314
Equity—agency	928	-0.354	1.924	-1.052	-0.996	-0.589	-0.096	0.000**	0.000**	0.053	0.256
Equity—identity	928	0.434	0.350	-0.031	-0.006	-0.105	-0.110*	0.338	0.994	0.019	0.470

Implementation measure	*n*	Covariates						Significance tests			ICC NSI
		Cohort		Outcome area				Wald test			
		Cohort 2	Cohort 3	8GOT	9GOT	CROT	Late entry	Cohort	Outcome area	Late entry *p*-value	
Equity—overall	928	0.435	0.976	-0.557	-0.432	-0.484	-0.090*	0.071	0.128	0.036	0.332

Note. 8GOT = 8th grade on track; 9GOT = 9th grade on track; CROT = College-ready on track; ICC = intraclass correlation; NSI = Networks for School Improvement

The reference category for cohort is Cohort 1. The reference category for outcome area is well-matched postsecondary.

The ICC for schools within NSI is not reported because it is near zero for all models. The remaining variation is between time points within schools.

n = the number of school observations in the model. Schools in the model had between one and three observations.

*$p < .05$. **$p < .01$.

Exhibit F2. Analysis of the Relationship Between CI Implementation Measures and Grant Year

Implementation measure	*n*	GY 5 coefficient	GY 5 *p*-value	Variance component				
				GY 5	Intercept	Covariance (GY 5, intercept)	School	Time points
At least one cycle initiated	799	-0.080	0.200	0.064	0.024	0.008	0.001	0.123
Number of cycles initiated	539	0.432	0.583	10.815	3.688	-4.429	0.609	1.940
Number of cycles completed	539	0.413	0.408	4.168	1.529	-0.625	0.501	1.746
Percentage of cycles completed	539	17.426	0.101	1926.279	713.197	-554.718	68.454	591.800
Change idea	539	0.092	0.574	0.421	0.228	-0.193	0.033	0.338
Use of structured tools	539	0.189	0.614	2.382	1.146	-1.034	0.006	0.807
Root cause analysis	920	0.447	0.006	0.604	2.238	0.191	1.648	0.203
Aim statement	920	0.306	0.541	5.773	5.184	-3.041	0.052	1.856
Theory of practice improvement	920	-0.261	0.445	2.747	2.569	-1.793	0.000	0.472
Data use	920	0.107	0.638	1.150	0.993	-0.567	0.045	0.574
Equity—access	920	0.726	0.023	2.435	0.500	-0.760	0.000	0.198
Equity—achievement	920	0.450	0.143	2.219	0.867	-0.794	0.000	0.315
Equity—agency	920	-0.233	0.213	0.825	0.600	-0.548	0.011	0.136
Equity—identity	920	-0.205	0.240	0.715	0.388	-0.221	0.013	0.121
Equity—overall	920	0.183	0.336	0.860	0.273	-0.284	0.002	0.084

Exhibit F3. Analysis of the Relationship Between Enabling Conditions and Disciplined Inquiry Cycles Measures

	At Least one cycle initiated		Number of cycles completed		Number of cycles initiated		Percentage of cycles completed		Change idea	
Enabling condition	β	*n*	β	*n*	β	*n*	β	*n*	β	*n*
Planning time	0.072*	679	-0.129	483	0.116	483	-2.259	483	0.018	483
CI team meeting frequency	0.002	657	0.222	472	0.341	472	0.226	472	0.148*	472
School supports	0.012	649	0.032	473	-0.010	473	0.697	473	0.010	473
School leader involvement	-0.010	744	-0.034	536	0.070	536	-1.900	536	0.048*	536
CI team stability—within year	0.149	742	0.203	542	-1.764*	542	12.960	542	0.178	542
CI team stability—across years	0.085	531	-0.055	370	-0.177	370	10.633	370	0.192	370
NSI support for knowledge management	0.010	651	0.214	472	-0.109	472	5.179	472	0.100	472
NSI supports to schools	0.014	650	0.351*	472	0.294	472	4.090	472	0.109	472
School climate—effective leader	-0.046	173	0.010	108	-0.052	108	-0.852	108	-0.125	108
School climate—collaborative teacher	0.011	157	-0.090	101	-0.049	101	0.621	101	0.018	483

Note. See Appendix E for full details on the enabling conditions analyses. The coefficient displayed here is for the enabling condition. NSI = Networks for School Improvement; CI = continuous improvement.

*$p < .05$. **$p < .01$.

Exhibit F4. Analysis of the Relationship Between Enabling Conditions and CI Core Parameters Measures

	Use of structured tools		Root cause analysis		Aim statement		Theory of practice improvement		Data use	
Enabling condition	**β**	***n***	**β**	***n***	**β**	***n***	**β**	***n***	**β**	***n***
Planning time	0.109	483	0.289**	773	-0.058	773	-0.109	773	0.148*	773
CI team meeting frequency	0.235	472	-0.437**	755	0.152	755	-0.014	755	0.015	755
School supports	0.050	473	0.089**	739	0.002	739	-0.001	739	0.030	739
School leader involvement	0.045	536	-0.039	844	0.013	844	-0.034	844	-0.012	844
CI team stability—within year	0.418	542	0.098	841	-0.436	841	-0.496*	841	0.026	841
CI Team Stability – across years	0.762**	370	-0.239	563	0.339	563	-0.088	563	0.251	563
NSI support for knowledge management	0.193	472	0.188	741	-0.015	741	-0.027	741	0.162*	741
NSI supports to schools	0.329**	472	0.177	742	0.090	742	-0.007	742	0.221**	742
School climate—effective leader	-0.044	108	0.001	227	-0.018	227	0.090	227	-0.048	227
School climate—collaborative teacher	0.109	483	0.289	773	0.069	211	0.091	211	0.006	211

Note. See Appendix E for full details on the enabling conditions analyses. The coefficient displayed here is for the enabling condition. NSI = Networks for School Improvement.

*$p < .05$. **$p < .01$.

Exhibit F5. Analysis of the Relationship Between Enabling Conditions and Equity in CI Measures

Enabling condition	Access		Achievement		Agency		Identity		Overall	
	B	*n*	β	*n*	β	*n*	β	*n*	β	*n*
Planning time	-0.109	618	-0.006	618	-0.062	618	0.032	618	-0.038	618
CI team meeting frequency	-0.080	600	-0.159	600	0.049	600	0.089	600	-0.025	600
School supports	0.004	590	0.010	590	0.000	590	-0.003	590	0.003	590
School leader involvement	-0.059*	685	-0.051	685	0.037*	685	0.020	685	-0.013	685
CI Team Stability—within year	0.093	804	-0.192	804	0.252*	804	0.139	804	0.076	804
CI team stability—across years	0.199	563	0.240*	563	0.093	563	0.005	563	0.134*	563
NSI support for knowledge management	-0.040	590	-0.046	590	0.037	590	-0.002	590	-0.015	590
NSI supports to schools	-0.108	591	-0.118	591	0.000	591	-0.019	591	-0.064	591
School climate—effective leader	0.022**	148	0.036	148	-0.030	148	0.005	148	0.008	148
School climate—collaborative teacher	0.099	141	0.073	141	0.010	141	0.042	141	0.054	141

Note. See Appendix E for full details on the enabling conditions analyses. The coefficient displayed here is for the enabling condition. NSI = Networks for School Improvement.

*$p < .05$. **$p < .01$.

Exhibit F6. Analysis of the Relationship Between CI Implementation Measures and CI Diffusion and Change in Climate

Implementation measure	Other staff adopt CI practices		Other staff adopt change ideas		School climate—effective leader		School climate—collaborative teacher	
	β	*n*	β	*n*	β	*n*	β	*n*
At least one cycle initiated	0.172*	647	0.166	233	-0.061	78	-0.261	85
Number of cycles initiated	0.001	470	0.035	173	0.108	66	0.071	68
Number of cycles completed	0.010	470	0.058*	173	0.042	66	0.027	68
Percentage of cycles completed	0.000	470	0.003*	173	0.003	66	0.002	68
Change idea	0.007	470	0.144*	173	0.030	66	-0.022	68
Use of structured tools	0.045	470	0.083*	173	0.101	66	0.075	68
Root cause analysis	0.054**	738	-0.002	239	-0.025	98	-0.095	105
Aim statement	0.026*	738	0.021	239	-0.054	98	-0.041	105
Theory of practice improvement	0.017	738	0.079*	239	-0.089	98	-0.086	105
Data use	0.130	738	0.100*	239	-0.063	98	-0.095	105
Equity—access	0.023	589	0.042	239	-0.195	98	-0.323	105
Equity—achievement	-0.024	589	0.001	239	0.086	98	-0.129	105
Equity—agency	0.004	589	0.000	239	-0.240	98	-0.360**	105
Equity—identity	-0.053	589	-0.014	239	0.093	98	-0.039	105
Equity—overall	-0.015	589	0.025	239	-0.110	98	-0.455*	105

Note. See Appendix E for full details on the CI diffusion and change in climate analyses. The coefficient displayed here is for the implementation measure.
*$p < .05$. **$p < .01$.

Exhibit F7. Results of Analysis of the Relationship Between CI Implementation Measures, Enabling Conditions, and Sustainability Measures

Implementation measure or Enabling condition	School continues to have a CI team		School continues to use CI approaches	
	β	*n*	β	*n*
At least one cycle initiated	-0.0386	114	-0.0986	74
Number of cycles initiated	-0.0180	93	0.0053	58
Number of cycles completed	-0.0531	93	0.0842***	58
Percentage of cycles completed	-0.0025	93	0.0022	58
Change idea	-0.0083	93	0.0865	58
Use of structured tools	-0.0527	93	0.0562	58
Root cause analysis	-0.0025	115	-0.0105	74
Aim statement	-0.0353	115	-0.0450	74
Theory of practice improvement	-0.0375	115	-0.0402	74
Data use	0.0493	115	-0.0292	74
Equity—access	0.0026	115	-0.0846	74
Equity—achievement	0.0065	115	0.1264	74
Equity—agency	0.0353	115	-0.2099**	74
Equity—identity	0.0445	115	0.0806	74
Planning time	0.2653*	107	-0.0253	70
CI team meeting frequency	-0.0556	106	0.0952	71
School supports	0.0125	125	0.0085	83
School leader involvement	0.0638*	136	-0.0420	87
CI team stability—within year	0.1774	111	-0.2319	71
CI team stability—across years	0.0654	97	0.0958	62
NSI support for knowledge management	-0.0237	126	0.0958	83

Implementation measure or Enabling condition	School continues to have a CI team		School continues to use CI approaches	
	β	*n*	β	*n*
NSI supports to schools	-0.0546	127	-0.1101	84

Note. See Appendix E for full details on the sustainability analyses. The coefficient displayed here is for the enabling condition. NSI = Networks for School Improvement. *N* of schools with data for this relationship equals less than 30—results omitted.

*$p < .05$. **$p < .01$.***$p < .001$

+

Appendix G. Survey and Interview Protocols

The survey and interview protocols for the Networks for School Improvement (NSI) school leader survey, intermediary interviews, climate surveys, and case study interviews are available upon request.

About the American Institutes for Research®

Established in 1946, the American Institutes for Research® (AIR®) is a nonpartisan, not-for-profit institution that conducts behavioral and social science research and delivers technical assistance both domestically and internationally in the areas of health, education, and the workforce. With headquarters in Arlington, Virginia, AIR has offices across the U.S. and abroad. For more information, visit www.air.org.

AIR® Headquarters
1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
+1.202.403.5000 | **AIR.ORG**

30058_04/26