

EDUCATION REFORM IN NEW YORK CITY

AMBITIOUS CHANGE IN THE NATION'S
MOST COMPLEX SCHOOL SYSTEM

Edited by Jennifer A. O'Day,
Catherine S. Bitter,
and Louis M. Gomez

Harvard Education Press
Cambridge, Massachusetts

Lessons for Other Jurisdictions

The problems addressed by the Children First reforms—including how and what to centralize and what to leave to the discretion of schools, how to ensure access for all children to high-quality teaching and opportunities for success, and how to motivate and sustain improvements over time—are issues common to all urban systems. Understanding the DOE's theory of action can pose alternatives for leaders elsewhere to consider, as well as specific tools that could be incorporated into other efforts. A few cautions are in order, however. The first is that New York City had a wealth of instructional capacity in the form of strong instructional leaders in many of the community school districts and support provider organizations on which to draw. Second, everyone with whom we spoke noted that the first phase focus on instructional coherence laid the necessary foundation for an empowerment approach. To the extent that other districts lack either the capacity for instructional support or school-level coherence, empowerment may not lead to improvements instructionally. Third, the DOE's accountability-based approach is unusually comprehensive in its generation and use of information on both leading and trailing indicators and in its structures to support professional collaboration. Piecemeal adoption of particular tools may not be strong enough to produce meaningful change. And finally, it is important to remember that many of the fundamental tenets of the reforms are as yet unproven. Scores and graduation rates have risen, but we do not know why, and we do not know whether this growth will be sustained over time.

6

Collaborative Inquiry to Expand Student Success in New York City Schools

Joan E. Talbert

Developing school capacity to continuously improve student achievement is essential to the success of New York City's strategy to empower schools and hold them accountable for results. The Inquiry initiative is the New York City Department of Education (DOE)'s approach to developing this capacity. Its goal is to develop school administrators' and teachers' skills in using multiple forms of student performance data to diagnose and close achievement gaps and to create school cultures in which educators collaborate in using evidence as the basis for instructional decisions to expand student success.

New York City's focus on building professional capacity for evidence-based practice stands in sharp contrast to conventional district approaches to improving student achievement. Most feature teacher professional development in content instruction and/or the implementation of curricula with fidelity. They promote particular standards for teacher performance in the classroom. An inquiry approach shifts the focus to student performance and calls on teacher teams to bring all students up to grade-level standards. Teachers are asked to diagnose the learning needs of struggling students and design instructional responses and system changes that meet their learning needs.

This chapter's three purposes are to:

- Describe the evolution of NYC's Inquiry initiative
- Illustrate how the inquiry model works to improve student achievement
- Point to challenges and dilemmas system leaders face in promoting and sustaining inquiry-based school reform

The description of NYC’s evolving inquiry design and resources draws on interviews and conversations with reform leaders inside and outside the system during 2006 to 2010. Illustration and analysis of school outcomes and implementation challenges draws on longitudinal research in NYC high schools.¹ The discussion of system challenges captures issues that have surfaced in NYC and in other districts pursuing inquiry-based school reform.²

Evolution of NYC Inquiry Initiative

NYC’s Inquiry initiative takes a continuous improvement frame on the problem of educational reform. In the long run, the system’s capacity to significantly improve student achievement depends on each school’s use of data to address all students’ learning needs. The problem of change, in this view, is that every school has a *sphere of success*—a group of students with whom it is currently successful. The challenge is to bring more and more students into this sphere.

The Inquiry initiative assumes that every school has a particular student population and pattern of skill gaps, as well as programs and policies that systematically limit the pool of successful students. It thus asks each school to analyze student performance data to determine why some students do not succeed—what skill gaps are not being addressed by the curriculum and how the instructional decision making systems limit success—and to respond effectively to accelerate their learning. This approach may seem straightforward. However, it challenges the conventional assumption that some students will fail to meet standards regardless of teacher efforts, and it brings into question a school’s established instructional culture. A design for inquiry therefore must be strategic in shifting teachers’ and administrators’ thinking about why students struggle and what can be done to ensure that they meet their potential.

The DOE modeled its original design for school inquiry after one created in a local administrator credentialing program called SAM (Scaffolded Apprenticeship Model), currently in its fifth iteration. Developed through a 2004–2005 pilot, SAM marries inquiry-based school reform with leadership development.³ The DOE piloted SAM’s design in over three hundred Empowerment Schools in 2006–2007 and launched the Children First Intensive (CFI) inquiry initiative systemwide in 2007–2008. In 2009–2010, the initiative was refined and renamed *collaborative inquiry*. Table 6-1 provides an overview of the evolving inquiry design and resources through 2009–2010.

The Inquiry Team Model

The inquiry team (IT) design features *teams* that collaborate to improve student learning; *tasks* of examining student work and data, identifying learning targets and instructional strategies, and using assessments to evaluate outcomes; and *tools* that scaffold

TABLE 6-1 Evolution of NYC’s inquiry initiative design and resources: 2006–2010

	2004–2010 SAM program (Baruch College—New Visions (NV); four iterations)	2006–2007 DOE’s ESI pilot (Empowerment Schools; N = 300)	2007–2008 CFI inquiry team initiative (all NYC schools)	2008–2009 CFI inquiry team initiative (all NYC schools)	2009–2010 CFI collaborative inquiry initiative (all NYC schools)
Design and resources					
Who’s involved?	Teacher team in certification program (involving twenty-eight schools, four in multiple iterations)	School team of principal and teachers	School team of principal, instructional leadership team (ILT) members, other staff; school data specialist	Two (or more) school teams of ILT, teachers, staff; school data specialist; two-year 90% goal identified	Teacher teams at grade-level and by content/SLC (90% goal); cross-cutting school team
Inquiry model	Team uses data to select up to fifteen target students and identify skill gaps and learning targets; designs and assesses instructional response to move students; makes system changes; spreads inquiry	SAM model	SAM model; target students from lowest-performing third of student population	SAM-DOE model expanded to target students who were not successful but not necessarily lowest third	Team examines student work/data and teacher work; identifies, implements, and evaluates an instructional strategy. Core team uses results to lead schoolwide change and capacity building
Inquiry tools and resources	Low-inference transcripts (LITs) of target students’ classes; coaching protocol; NV data platforms	SAM tools	Inquiry target tool; periodic assessments; ARIS tools; progress report; quality review (QR); CFI Sharepoint site (website containing training materials shared between SAFs and SSOs); Inquiry Team interface (early version of inquiry spaces)	Handbook with SAM tools and guidance; same tools and resources as 2007–2008.	Same, with revised inquiry diagram and QR rubric; collaborative inquiry resources and inquiry spaces (online work spaces located within ARIS)

continued

TABLE 6-1 continued

	2004-2010 SAM program (Baruch College—New Visions (NV); four iterations)	2006-2007 DOE's ESI pilot (Empowerment Schools; N = 300)	2007-2008 CFI inquiry team initiative (all NYC schools)	2008-2009 CFI inquiry team initiative (all NYC schools)	2009-2010 CFI collaborative inquiry initiative (all NYC schools)
Design and resources					
On-site resources	Tuition support; school funds for dedicated inquiry time	Per session allocation for team members to meet before and after school	Per session allocation for team members to meet before and after school	Reduced monetary allocation to support inquiry work	Reduced monetary allocation to support inquiry work
Inquiry team support/training	Weekly seminars and coaching with SAM facilitator; weekly facilitator training	Support from ESI network leader and achievement coaches; monthly network leader training Two three-hour workshops for principals preparing for the introduction of inquiry work and roll-out of the accountability tools	Support from network, SAF; monthly SAF/SSO training; monthly principal network meetings; monthly data specialists training	Same support; SSOs designed training for network teams, and SATIFs Stipends for principal and teacher leaders to support the inquiry professional development in each school networks	Support from principal/team leaders, citywide monthly cluster institutes Support from network, SATIF, achievement coaches; network sharing Stipends as per 2009

the inquiry work. The model was designed to develop teachers' skills in using student work and assessment data to analyze and address particular skill gaps, thus challenging assumptions that not all students can succeed. It empowers them to make decisions about what and how they are teaching and, in turn, to influence schoolwide decisions.

The CFI model replicated SAM's design. However, it was implemented without the credentialing program's rigorous assignments, weekly seminars, and ongoing feedback and support from an instructor well prepared to facilitate inquiry-based school reform. The IT model did include support from a senior achievement facilitator (SAF) who was an experienced educator/administrator trained to lead inquiry, as well as an evolving set of assessments and online tools to support teams' inquiry work.

Inquiry Team. During the first year of implementation (2007-2008), a school inquiry team comprised the principal and several teachers and staff who represented a broad range of expertise. IT members were responsible for collaborating to use data to improve the success of struggling students. In addition, they were expected to lead colleagues to use instructional responses effective with struggling students, to identify and improve instructional decision making systems, and to spread inquiry practices in the school.

Targets. The model emphasizes "getting small in order to go big" with inquiry-based improvement.⁴ This means focusing on a small group of target students and a specific learning target as a starting place for school reform. DOE guidelines for practice instruct the IT to first identify a content area that most needs improvement according to the data and then to identify a target population of struggling students (the lowest-performing third) and select a small group of twelve to fifteen students as a focus for their inquiry. The team then is to use multiple assessments to identify a skill gap prevalent among the target group, such as reading, and then to home in on a subskill (e.g., comprehension) and a particular learning target relevant to that gap (e.g., topic recognition or using context cues).

The approach of beginning inquiry with small learning targets emerged from the early experience of SAM facilitators. School teams were overwhelmed by the need to analyze large amounts of data and translate the data into ways of helping students meet grade-level standards. The facilitators responded by focusing the work on specific targets for intervention. They reasoned that this made the work manageable in scope and potential for improving students on a given skill. Absent a push to stay small, some teams gave up on what they perceived as the need to bridge very large skill gaps, doubting that they could make a difference.

Tasks. The IT model specifies three broad inquiry phases imported from the original SAM model. As described in the DOE's 2008 *Inquiry Handbook*, the phases and steps are:

- *Phase I. Identify Focus, Students, and Targets:* Use data to identify a focus area in which the school is not doing well (e.g., ELA or mathematics), identify a

schoolwide target population of students with which the school has not been successful historically, define a small group of target students, set a long-term goal, define learning targets and short-term goals.

- *Phase II. Move the Students:* Analyze conditions of learning for the target students, design and implement an instructional change strategy, evaluate and revise based on interim progress measures.
- *Phase III. Move the System:* Analyze instructional decision making systems that produce conditions of learning, design and implement a system change, evaluate and revise based on interim progress measures.

The tasks are meant to guide the work of a school team and to provide a pathway for continuously bringing more students into the sphere of success. Both SAM and the DOE model have moved incrementally toward spreading the inquiry model to multiple teams across a school.

Tools. Resources and tools designed to support school teams' inquiry include data platforms with results of periodic assessments, formative assessments and protocol for their development, and protocol for looking at student work. The DOE has developed an arsenal of tools over the course of the Inquiry initiative, many of which are housed in ARIS (see table 6-1).⁵

One tool developed through SAM and built into the IT process is the *low-inference transcript* (LIT) in the classrooms of target students. The LIT is a means of documenting the class experience of students by writing a verbatim script of classroom talk and activity. It discourages observers from making inferences about "quality of instruction," or promulgating the idea that a student is "lazy." It is meant to provide valuable data for understanding the school curriculum as taught and experienced by students—rather than as it exists in a plan or in teachers' minds. LITs from multiple classrooms allow a team to identify cross-class trends that help to inform their next steps. As discussed in the next section, inquiry teams have found this to be a powerful resource for their work because it shifts their focus from instructional delivery to student learning.

Launching the System's Inquiry Team Initiative

The DOE asked all NYC schools to implement the IT model in school year 2007–2008. Every school was required to establish an IT that included the principal and key school leaders and staff. The team was to designate a data specialist who would attend monthly meetings to learn how to use NYC's data systems and to share effective practices with other schools, as well as to lead data analysis in the school team.

Two major and concurrent system changes compounded the challenge of getting the systemwide IT initiative off the ground and running—restructuring from regional authorities to School Support Organizations (SSOs) and networks, and the initiation of the Progress Report and Quality Reviews by the Office of Accountabil-

ity.⁶ As a consequence, the DOE faced substantial resistance as it pushed to develop systemwide capacity for inquiry-based school improvement.

Resistance. Resistance came from the perceptions by principals and some SSO leaders that the IT initiative was a top-down mandate for change in school practice. Principals were reeling from major changes in system organization and accountability demands at the time they were asked to launch the IT initiative. A DOE leader commented: "We found a lot of angry people who . . . were very upset that Regions [regional authorities] were going away after all the buildup of the Regions and all of that work. They just didn't understand why this was happening."

Given their role as service organizations, some SSOs shied away from pushing the implementation of IT practices, fearing that school leaders would see them as an arm of the Office of Accountability. As a result, as one DOE leader put it: "We had different, mixed messages being sent out to the network leaders . . . In some cases [when the Senior Achievement Facilitator was a strong inquiry facilitator], network leaders were very supportive. In other cases, not . . . So it was an interesting first year."

In turn, school leaders received varying messages and levels of support for developing an effective inquiry team. Many principals did not understand how inquiry would help them increase student achievement and complied only minimally with the IT requirement. SSO leaders, SAFs, and network leaders could help schools make connections between inquiry and student outcomes only if they understood and bought into the model themselves. Because the inquiry initiative pushed against more conventional approaches to instructional improvement, such as teacher professional development in content instruction, many of those charged with supporting the inquiry work were not on board or lacked the skills to push the reform, or both.

Capacity Building. Although the DOE was developing a robust data system to support school inquiry, the system was not up and running when the inquiry initiative was launched. To compensate, the DOE developed a tool that was essentially an Excel file of student data that was not very user friendly.

Training was a major challenge during this phase of the initiative. In contrast to schools' voluntary participation in SAM and the DOE pilot in Empowerment Schools, the rest of NYC schools had not opted into the inquiry team initiative or its training. Moreover, a top-down approach to school improvement ran counter to the rhetoric of school autonomy and accountability.

In addition, the inquiry model included many facets and phases. A DOE leader commented that trainers were challenged "to make the work less theoretical and embed it in the actual work that people are engaged in . . . so that it becomes *practical* and they see the connections . . . That's always the biggest challenge."

The range and diversity of school readiness for inquiry posed yet more training challenges. A key factor was the principal's level of comfort with distributed leadership and willingness to collaborate and share decision making with teachers. A further readiness issue was the fit between inquiry practice and the school's culture.

According to one DOE leader, “Some schools were in crisis. Other schools had their own things in place and were moving very nicely and felt this was now a mandate—like ‘Who needs this? I’m doing fine with my school.’”

Nevertheless, some schools embraced the IT model and took advantage of DOE resources designed to support development of inquiry practices. According to an external evaluation of the first phase of the CFI inquiry initiative, 54 percent of school teams had completed a full inquiry cycle by the end of 2007–2008 and 75 percent by the early months of the 2008–2009 school year.⁷ (It should be noted, however, that these statistics do not distinguish between ritual inquiry practice and deep inquiry cycles that boost target students’ performance. Research on inquiry-based reform in NYC and elsewhere points to a qualitative difference between implementing surface features and embracing principles and “stance” of inquiry to improve student achievement.)

Implementation Challenges. The DOE used feedback from network and school leaders, along with its internal evaluation, to identify three major implementation challenges that shaped the 2009–2010 revisions.

First, although the DOE considered the Inquiry Team model relevant and valuable for all schools, some principals perceived it as out of sync with their school culture and implemented it in name only. For example, educators in some small schools that embrace a whole child philosophy of education regarded the use of fine-grained data to diagnose and address student skill gaps as inconsistent with their beliefs about how to improve student achievement. In such cases, the principal rejected data-based inquiry as a model for school improvement. Yet school empowerment and accountability hinge on leadership by principals, and their buy-in to the inquiry model was crucial to its success as an engine for school reform.

Second, it appeared to district leaders that some teams were spending too much time on data analysis. Many never moved to making fine-grained evidence-based inferences, formulating hypothesis, and testing those hypotheses in practice. One DOE leader framed the problem in these terms:

We saw that as teams got together, as they began to look at the data, they spent a lot of time looking at the data trying to identify their target population or what they wanted to do—their learning target—and they didn’t get to the real work. And so we had to put some deadlines in, at least some benchmarks—“You should be at this stage”—and [try] to prod them along and push them along . . . and [work] very hard with the Design Team [SAFs and SSO leaders] to help them understand inquiry and the entire process and then help them facilitate the teams in moving along the process. And that was really difficult the first year.

Some of those teams lacked sufficient guidance for assessing student skill gaps and designing effective instructional responses, and may have needed more time and support for moving to an action phase that could make a difference for struggling students.

Third, among teams that *did* design and implement an instructional response, many created responses that may have made a difference for individual students but

did not impact classroom instruction (e.g., tutoring, after-school sessions, Saturday classes). Some in the DOE viewed this as not meeting the goals of the inquiry initiative. Yet SAM facilitators came to regard such responses as “trying out” or piloting potential schoolwide interventions and found that teams sometimes needed support in applying what they learned in out-of-classroom settings to the instructional core.

The DOE revised its inquiry initiative in 2009–2010 to address these challenges, with changes designed to make the model more manageable and embedded in instruction. One DOE administrator explained the thinking this way:

You actually need to simplify the task significantly in order for it to take root and spread . . . When [we] boiled down what the intention of that original model was, it involved looking at student work and student data and looking at the corresponding teacher work and practice and developing a theory about why some portion of the kids are not succeeding in that environment, and a strategy to help them succeed that you then monitor through the use of common assessment tools. And with that feedback loop, make adjustments to some key levers that are at your disposal . . . What’s being taught, how it’s being taught, how you’re assessing what’s being taught[, and] how adults are learning in the school . . .

You have to start at the point where people are getting into the habit of looking at student work seriously . . . And part of the habit we’re trying to develop is a way of thinking and a cultural shift. So it actually doesn’t have to be perfect. Like it’s okay for people to be focusing at a more generic [skill] level if the habit is actually forming. Because . . . if you get good at this, you do get driven down to the more specific.”

The revised inquiry model and guidance for school implementation were intended to spread and deepen inquiry within NYC schools.

Refining the School Inquiry Model

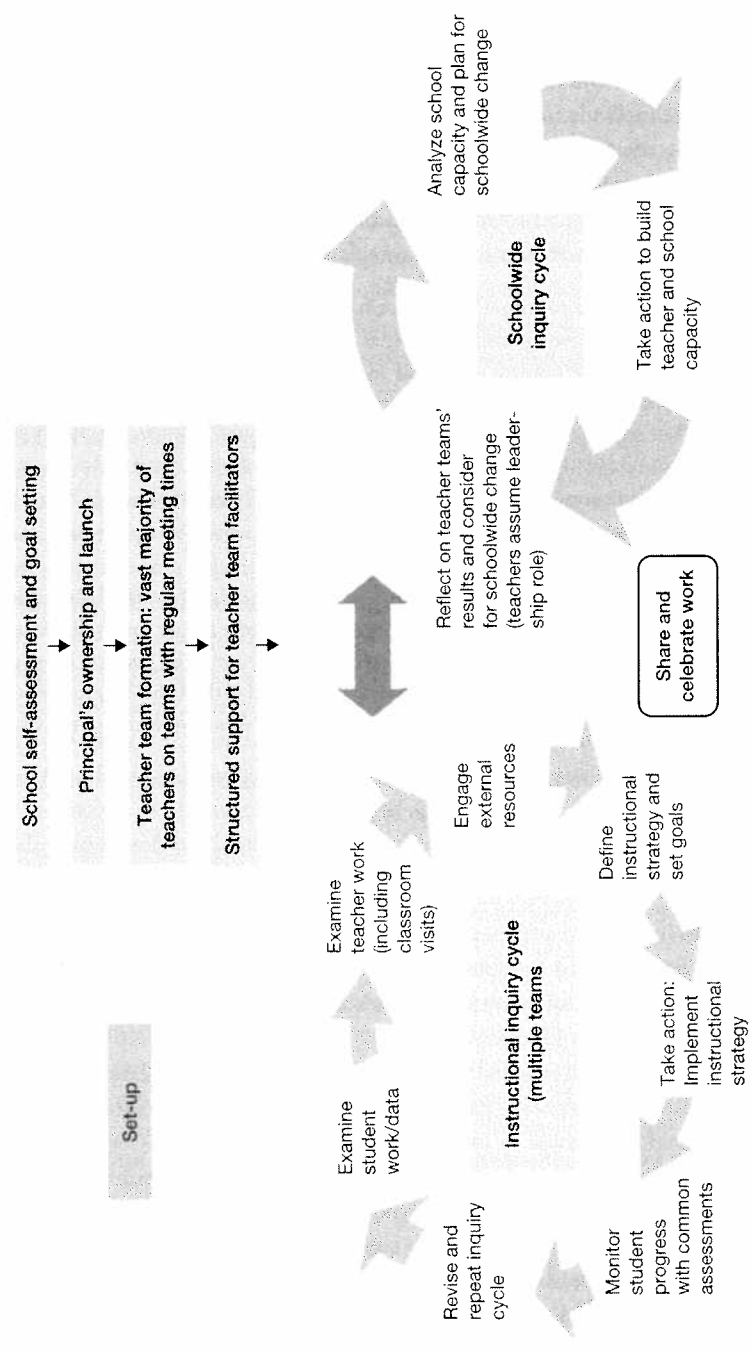
The DOE introduced a refined *collaborative inquiry* model in the 2009–2010 school year (see figure 6-1 for a graphic of its current design).

Refinements call for quantitative and qualitative changes in schools’ inquiry practices from the earlier IT model.

First, each school is asked to involve most teachers in collaborating on inquiry with a team of colleagues—for example, a grade-level team, a subject department or course group, or a house or small learning community (SLC) in restructured high schools.⁸ School administrators are called on to establish inquiry teams across the school, schedule their common meeting time, and designate and support a teacher facilitator for each team. This change seeks to spread inquiry across the school and to expand teacher leadership.

Second, the model specifies that the principal’s goals for school improvement are the starting place for teacher teams’ inquiry. A set-up phase asks principals to lead a schoolwide self-assessment process and then galvanize staff toward common work for the year. The focus for common work is to be informed by data and information

FIGURE 6-1 Collaborative Inquiry process (DOE, 2009-1010)



previously gathered and examined as areas of need. This change responds to principals' push-back on the original design and seeks to leverage their buy-in and support for inquiry-based reform.

Third, guidelines for inquiry practice place greater emphasis on classroom instruction than the original inquiry model. Teacher inquiry teams assess student performance against grade-level standards in a content area and in relation to their instruction. The model specifies these steps in the cycle: examine student work/data, examine teacher work (including classroom visits), engage external resources, define instructional strategy and set goals, implement instructional strategy, monitor student progress with common assessments, and revise and repeat the inquiry cycle (see figure 6-1). It continues the original guidelines for focusing on a small group of target students not meeting their potential and staying small to address particular learning targets using an instructional strategy. Because inquiry is conducted by grade-level teacher teams, the work is essentially embedded in instruction. This redesign is intended to deepen inquiry and bring it into classroom practice. Principals and the inquiry team leaders they designate are expected to facilitate the development of inquiry practice and norms.

Consistent with these changes in inquiry guidelines, the DOE revised its Quality Review (QR)—the primary tool for giving schools feedback on their progress toward an inquiry culture (see Childress et al., chapter 4 in this volume). Quality Review ratings on multiple dimensions of school culture and practice provide evidence of whether a school is developing capacity for improvement, potentially counterbalancing a weak Progress Report. Revisions changed language and scoring from what some system leaders regarded as “narrow, quantitative” criteria (e.g., the school has at least four teams doing X focus on inquiry) toward a more descriptive rubric (e.g., teachers are working to improve their X instruction). This aimed to avoid the tendency of some schools to jump through hoops to meet quantitative measures, as well as to prompt a more holistic assessment of the school’s instructional and professional culture.

A DOE staff member explained changes in the QR in these terms: “We’ve spent three years building a data culture. And this tool, the Quality Review, has been a leverage point [along with the Progress Report]. It seemed that the time had come to push on making instruction and instructional coherence as sort of our organizational program, to really be the point of the Quality Review. [We’re] no longer building a data culture of the school.”

Training demands for implementing the new rubric during 2009–2010 were considerable, particularly since this was the first year that external reviewers were not conducting the QRs. The DOE provided monthly training sessions for a new cadre of internal reviewers, as well as network leaders who opted to attend, to “build a deeper understanding of the quality that we’ve defined in the rubric with its twenty indicators.” The learning curve had to be steep, since the DOE conducted five hundred QRs during 2009–2010.

New Challenges

Extending collaborative inquiry to all teachers in NYC schools posed new challenges for system leadership. Principals and network leaders are responsible for leading schoolwide inquiry, yet some lack the commitment, understanding, and skills to do so. In some schools, the original IT design developed a cadre of inquiry leaders, but this was not true across the board. DOE leaders are challenged to address increasing inequalities in inquiry leadership capacity at all system levels so that all students get the benefits of teachers' collaborative inquiry.

Developing broad inquiry leadership at the school level is critical to the initiative's success. Research and practice testify to the fact that principals and team facilitators play key roles in leveraging and supporting school culture change.⁹ In leading change, they prompt teachers to rethink beliefs and assumptions about students, colleagues, and their own abilities that constrain progress on inquiry to expand student success.

In this new phase of inquiry-based reform, system leaders aim to develop a critical mass of teachers in each school who can facilitate the work of their teams in ways that bring about culture shifts to improve student learning. As one DOE administrator commented: "You can put a structure in place, but if there isn't a facilitator in that team that's going to push and keep it focused—and if that person doesn't have a place to reflect and process to get support themselves—it's much harder to make it successful."

Some networks have made strides in training teacher leaders to facilitate collaborative inquiry with their colleagues. The DOE is challenged to develop all networks' capacity to play this role in developing school-based inquiry leadership.

NYC's Inquiry initiative is intended to transform schools' professional culture toward internal accountability for continuous improvement. As a DOE administrator put it: "The inquiry team structure isn't just intended to facilitate getting good at the habit of inquiry. It's also intended to facilitate the breaking down of isolation between teachers, developing teacher leadership, and accelerating the spread of effective practice within the schools."

Our research in NYC high schools has addressed the questions of whether and how school professional culture shifts in schools that implement the inquiry model, as well as what conditions affect implementation. We find that the inquiry model works, when well implemented, to change teacher beliefs and practices in ways that improve student outcomes. However, the challenges of implementing the model are considerable, particularly in high schools, and some teams have lacked essential resources for change.

How Inquiry Works and Implementation Challenges

New York City high schools have implemented collaborative inquiry to widely differing degrees. At one extreme are a small number of schools with nearly five years of

experience developing robust inquiry practices through the SAM credentialing program. At the other end of the spectrum are schools that have lacked administrator and/or network support for implementing the inquiry model, and that only ritually carry out inquiry in team meetings, if at all. In between are the majority of schools that have been making steady progress, in good faith and with principal and facilitator support, over the first three years of the DOE's inquiry initiative.

Variation in schools' implementation of inquiry allows us to investigate the questions:

- Does the model work to change school culture and expand student outcomes when it is well implemented?
- What challenges do schools and teachers face in implementing the model and what resources make a difference?

Our research bearing on these issues includes (1) a two-year study of fourteen high schools involved in SAM's second iteration (2005–2007), including seven Autonomy Zone/Empowerment Schools, and (2) a subsequent ongoing study of inquiry work in over seventy New Visions Partnership Support Organization (PSO) schools, most with no experience in the SAM program. Included in both studies are four schools that have participated in SAM continuously beginning in 2005–2006 (dubbed *mature inquiry* schools for their work on implementing the model beyond the three years of the system initiative).¹⁰

To address the first question, we use five-year longitudinal case studies of the four mature inquiry schools. Evidence from their track record offers existence proof of the theory of change and predicts that inquiry-based reform will pay off in the long run for most NYC schools. Each of these schools developed an inquiry culture, shifted instructional perspectives and practices to address the needs of struggling students, and brought more students into the sphere of success. These schools represent an intersection of strong inquiry support and challenging school contexts. On one hand, teacher teams from these schools received intensive support in implementing the NYC inquiry model through the SAM program; on the other hand they are high schools—two large and two small—that present special problems for change (see Siskin, chapter 8 in this volume).¹¹

To address the second question, we use two years of survey data for all New Visions schools and case studies of twelve schools that represent contrasts in SAM experience, high school size, and grade level (most schools in the New Visions PSO are high schools, but we purposely included two elementary schools). Quantitative analysis identifies predictors of school progress in developing an inquiry culture. Qualitative case study data point to challenges that cut across schools that differ in their experience with inquiry and in school organization.

These broad and in-depth analyses of school inquiry practices and culture change offer insight into the developmental arc and phases, as well as hurdles, entailed

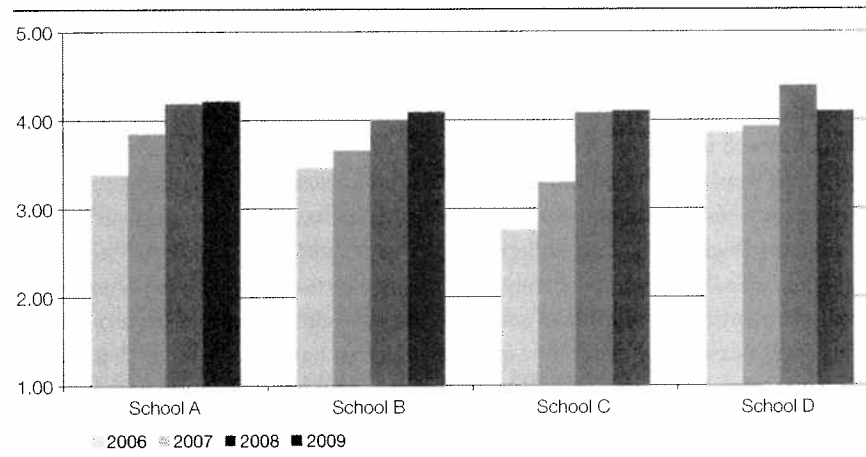
in developing a school culture of continuous improvement. Most fundamental is whether a team makes the qualitative shift from ritual enactment of surface activities and so-called single-loop learning to authentic collaborative practice and double-loop learning, in which members share habits of mind and practices of using evidence to diagnose and address student learning needs through instruction and system changes.

All school inquiry teams encountered technical, organizational, and cultural challenges in implementing the model for data-based decision making. Managing them depended on a skilled facilitator, assessment protocol, and principal commitment to inquiry-based reform and teacher leadership.

Inquiry Changes School Culture, with Tipping Point in Third Year

Annual faculty surveys in the four schools involved in SAM since 2005–2006 document the schools' steady progress toward a culture of inquiry-based improvement. Each school shows step increases on a measure of "culture of assessment use" for the first three years, followed by a plateau (see figure 6-2).¹² In the two large high schools organized into SLCs, SAM graduates lead inquiry with colleagues in their SLC, and the survey data capture development of inquiry cultures within and across these units. In the two small high schools, a school team including several SAM graduates leads inquiry with grade-level and content-area colleagues. Over time, increasing proportions of teachers were collaborating on using multiple assessments to focus instruction for struggling students. Much of this work focused on student skill gaps in literacy, such as identifying the main idea in a text or knowing academic vocabu-

FIGURE 6-2 School culture trends for mature inquiry schools: Teacher reports on assessment use



lary in or across a subject area. Teacher teams shared with school colleagues instructional responses for which they had evidence of success with their target students.

These schools' inquiry practices were a radical departure from the past. Three of the four schools began SAM with weak assessment cultures (note 2006 levels of assessment use for schools A, B, and C in figure 6-2). Their data use was limited mainly to reviewing standardized test results, particularly scores on Regents exams. Each teacher and subject department had considerable latitude in deciding whether and how to use finer-grained interim assessments. The 2006 baseline data show that teachers were about as likely to disagree as to agree that assessments were being used to inform instruction. The fourth school (school D) had a tradition of assessing individual student performance through portfolios submitted twice a year, so teachers' initial ratings of their assessment use were relatively high. Through SAM, however, the school made a qualitative shift toward using fine-grained skill assessments to identify and home in on learning targets for their struggling students, primarily recent immigrant English learners with weak academic preparation.

Survey results for spring 2009 show a plateau on the measure of assessment use after three years of change. New inquiry norms and practices were being sustained, and the schools were deepening their inquiry work in ways not well captured by the assessment use survey measure.

Further, in each school we observed particular kinds of culture shifts that accompanied teachers' developing inquiry practice. They illuminate both challenges for change and outcomes of inquiry:

- *Shared accountability:* As teachers worked in teams to diagnose and respond to specific learning needs of struggling students, they began sharing responsibility for the success of all students. Teachers moved from thinking about "my" students to "our" students, as well as shifting their attention from successful students to struggling students.
- *Norm of evidence-based practice:* Faculties developed the habit of using evidence of student performance to evaluate and improve instructional decision-making systems. Teachers moved from (1) relying on their intuition and past practice to using data to drive their instructional decisions and evaluate student learning, and (2) using summative assessments to measure student outcomes to using formative assessments to diagnose student learning needs.
- *Distributed leadership:* As teachers began taking leadership roles in their inquiry teams, ideas and norms about school leadership shifted from administrator decision authority and prerogative to widespread agency and responsibility for improving student success. Teacher teams became leaders of inquiry-based decision making for school improvement.

These fundamental shifts in professional norms and practices established conditions for sustainable inquiry-based improvement in the school.¹³

As expected, these mature inquiry schools evidence greater success in bringing struggling students into the sphere of success. Using 2008 “on track”/“off track” data for students in the 2009 graduation cohort who had scored below proficient in ELA in eighth grade—students who were struggling academically before entering the high school—we found that a significantly higher percentage are “on track for graduation or college readiness” in the four mature SAM schools than in non-SAM schools in the New Visions PSO (68 percent versus 34 percent). Further evidence comes from regression analyses that estimate the effect of a school’s inquiry implementation on student outcomes. We found that a school’s mean on the “culture of assessment use” survey measure predicts the percentage of its 2009 cohort students who are on track in 2008, with controls for the percentage of those students who entered with below-proficient eighth grade ELA scores.¹⁴

To what extent might these strong positive results be due to the fact that the four mature inquiry schools were early adopters? We know from decades of research that early adopters do better with any kind of innovation because their decision to adopt an innovation signals motivation and readiness to engage it. Perhaps the schools were already on their way to inquiry-based reform, and supports from SAM and the DOE initiative made little difference. Nonetheless, they offer existence proof that inquiry can be a vehicle for school culture change and improved student outcomes. They moved significantly beyond our baseline measures of their inquiry practices. The facts that teacher teams had intensive support from SAM facilitators and the principal endorsed their work point to resources that make a difference for school progress on the initiative.

Significant improvement in student outcomes reflects the gradual shifts in school culture we documented. These in turn reflect the shifts in teachers’ beliefs and perspectives on struggling students and classroom practices that come about through inquiry practice.

Inquiry Changes Teachers’ Classroom Practices

The substantive work of collaborative inquiry and leadership is unique to each school and team. Each encounters particular skill gaps of struggling students and facets of their school’s instructional culture and systems that keep the students outside the “sphere of success.” Yet just as collaborative inquiry engenders certain kinds of school culture changes across diverse schools, so too does it prompt particular shifts in teacher perspectives and classroom practices. Our research suggests that some shifts occur within the first year of implementing the inquiry model:

- *Shift in focus from teaching to student learning:* Most teachers in the inquiry teams we studied said that they had made a big shift in their thinking about classroom instruction. In their own classroom and in observing others, teachers’ focus moved from how the curriculum is being taught to what stu-

dents are learning. They experienced this shift as an important benchmark in the development of their inquiry skills and teaching practice. Many told us that doing LITs in their target students’ classrooms prompted this change in perspective. The tool prompted them to see instruction through the lens of struggling students. They learned that their ideas about “high-quality” teaching did not always mesh with struggling students’ learning needs. Teachers became aware that students had often missed critical segments of content instruction that state standards prescribed for earlier grades and that this content was not being offered to them in high school courses geared to grade-level standards.

- *Shift from summative to formative assessments of student learning:* Teachers moved from testing for grading purposes to using formative assessments to diagnose student learning needs and develop an instructional response. Going small in assessments to identify misconceptions and gaps in student understanding helped them create responses that accelerated the learning of struggling students. Further, teachers moved to better scaffold learning objectives for their lessons and ask students to give them information about their learning and struggles with particular content.
- *Shift from external attributions of student failure to instructional efficacy:* Teachers stopped perceiving student failure as something beyond their control. Explanations shifted from “miserable family circumstances” or “personal troubles” to skill gaps resulting from prior and current academic experiences. Addressing the gaps became the main concern. As teacher teams designed effective responses and saw the academic gains students were making, they developed a sense of instructional efficacy that carried over into their classrooms.
- *Shift toward on-demand professional development in content instruction:* In some schools, inquiry teams converged in their efforts to address skill gaps prevalent among struggling students, prompting a schoolwide instructional response. For example, after three years of SLC-based inquiry work, team leaders across a large high school reached consensus that student writing was a high-leverage skill domain. As a consequence teachers were eager for professional development (PD) to support their instructional responses. The principal brokered a series of on-site PD days with a literacy/writing expert whose work was enthusiastically received. This teacher learning agenda grew out of their diagnosis of student learning needs, rather than from the judgments of administrators about what teachers needed to know. Interestingly, as a baseline, the same literacy expert had come to the school several years earlier (before inquiry had shown the need for this kind of PD) and, by all accounts, teachers paid little attention. Demand for PD generated through inquiry into student learning needs made all the difference in teachers’ readiness to

learn and to make changes in their classroom practice. Teachers came to the PD eager to learn from the expert and eager to try out new practices in their classroom.

Such shifts in teacher perspectives and instructional practices interact with changes in school culture. For example, developing team norms of shared accountability for using inquiry to meet the needs of struggling students helped individual teachers shift their focus from delivering curriculum to diagnosing students' learning needs, while individual experiences of making a difference for struggling students helped tip the school toward an inquiry culture.

School Change Entails Technical, Organizational, and Cultural Challenges

We find that progress on data-based inquiry is not linear. Rather, it is bumpy and cyclical. As teachers move outside their comfort zone to develop new assessment and instructional practices, they grapple with the tug of old habits and mind-sets. Teachers report moving two steps forward and one step back, needing to relearn new practices and perspectives. They experience an "Aha!" only to encounter a new challenge. Some teams get stymied by the roadblocks they encounter and never get beyond superficial routines of data use; others become highly skilled in using data to continually improve student learning and success.

The resources a team can draw on for tackling the technical, organizational, and cultural challenges for change matter a great deal. Table 6-2 summarizes the challenges and resources that have made a difference in teams' progress on inquiry-based school improvement.

Technical Challenges. Schools began their inquiry work, whether through SAM or through the DOE's Inquiry initiative, with little prior experience in using student assessment data to design and evaluate their instruction. Most teams struggled to use multiple indicators of student performance, use assessment data and student work to identify prevalent skill gaps, and develop and use formative assessments to evaluate the success of an instructional response.

A team's ability to get up and running on inquiry cycles depended on having an assessment-savvy person to lead the work. The designated IT data specialist was a key resource in many schools. Through their monthly network meetings, these specialists learned the ins and outs of the DOE data system and how to analyze periodic assessment to identify specific skill gaps in student performance. Networking with colleagues from other schools also pointed to effective ways of leading school teams and innovative ways of organizing data.

Yet all teams struggled with the push to go small and identify a specific, manageable learning target that they could teach to and use to improve their instructional decision making. Not only did they need skills in looking closely at assessment data and student work but, to many teachers and administrators, the idea of going small to make a big difference was counterintuitive.

TABLE 6-2 Using inquiry to improve student achievement: Technical, organizational, and social-cultural challenges and resources

<i>Challenge for change</i>	<i>Facilitating conditions and resources</i>
Technical: Developing inquiry practice	
Using available system data to identify students outside the "sphere of success" and formative assessments to identify skill gaps	Assessment-savvy person on team Data system and summaries that include multiple measures and fine-grained data Data specialist meetings and network
Getting small: Focusing on target group of struggling students and homing in on a learning target for instructional response	Skilled facilitator to address resistance and keep the work focused
Assessing student learning outcomes of instructional responses in order to refine them	Skilled facilitator to guide development of pre/post assessment and provide or point to resources for instructional response
Organizational: Developing leadership	
Creating and protecting time for collaboration on inquiry	Principal commitment and priority for collaborative inquiry
Distributing leadership and developing teachers' capacity to lead inquiry in teams and the school	Principal delegation of authority to teacher leaders and inquiry teams
Social-cultural: Challenging constraining beliefs and habits	
Developing shared responsibility for student success	Administrator focus on results by teacher team; a press for team success
Shifting focus from teaching and curriculum delivery to student learning and skill gaps	Low-inference transcripts (LITs); administrator assurance that LITs are not for teacher evaluations
Shifting teachers' attribution of student failure away from external factors; developing their sense of instructional efficacy	Evidence of team success in accelerating student achievement; team presentations to colleagues and impact on schoolwide decisions

In some schools, the external facilitator (SAF or SAM instructor) helped the team get past frustrations of learning to implement the inquiry model. As one teacher put it: "The process was so frustrating at times that I think if there wasn't an outsider pushing you, we just would have said: 'No. It's not working.' Or, 'These are just the types of kids we get. And we're not going to be able to move them.' Just having an outsider to keep pushing you and still *be* there was critical."

Organizational Challenges. An inquiry team needs regular dedicated time for its work. Yet site administrators manage competing priorities for teachers' time and work outside the classroom, and schools vary widely in both frequency and reliability of time designated for teacher inquiry. Some teams floundered because their scheduled meeting time was often co-opted for another purpose, such as planning for summer school or professional development for a curriculum project. Absent a school priority for collaborative inquiry and protection of the schedule, the work stalls and teachers see it as a DOE mandate and take a compliance mentality.

In order to support collaborative inquiry, principals need to understand the *principles* and believe that it is an effective vehicle for instructional improvement. In schools where the principal was on board and strategic in involving teacher leaders, teams became effective. Using two years of data from our annual survey of IT members in New Visions PSO schools, we found a positive statistical effect of “principal support” on growth in “team functioning.”¹⁵ This does not imply that the primary change agent was the principal. Principals rated high on the inquiry team support scale included those who delegated leadership almost entirely to teacher leaders on the team. Indeed, broadly distributed leadership is fundamental to inquiry-based reform. A principal’s willingness to share decision authority is essential if collaborative inquiry is to take root.

Professional Culture Challenges. Reform leaders face opposition and constraints on change that stem from long-standing norms in teaching. Yet schools varied in how extensive and ingrained traditional norms were, since their reform histories or founding cultures may have pulled teachers away from conventions. Nevertheless, all schools are challenged to address constraining professional beliefs and habits.

The Inquiry initiative’s call for *teacher collaboration and shared accountability* for student success pushes against norms of privacy and individual responsibility for classroom instruction. The inquiry model’s clear focus on students appears to be a useful vehicle for building trust and “deprivatizing” classroom practice. Facilitators who made a difference held this focus when teachers retreated into privacy. Administrators pushed for teachers’ shared accountability by focusing on grade-level, department, and SLC performance in reviewing school progress.

Shifting teacher focus from curriculum delivery to student learning requires a reframing of high-quality instruction to focus on outcomes for struggling students. As noted, teacher teams point to the practice of LITs in classrooms of target students as a key lever for change. For example, after diagnosing target students’ gap in academic vocabulary, a team was taken aback to see in all its LITs teachers’ frequent use of terms such as *summarize*, *analyze*, *synthesize*, and *interpret*—realizing that the students could not comprehend such directions or access instructional content. School administrators and team facilitators play important roles in implementing this tool. Success depends on administrators making clear that the classroom observations are not being used for purposes of teacher evaluations. Facilitators support the developing teachers’ skill of scripting classrooms verbatim so that student experiences become accessible and available for developing instructional responses. The use of protocols to analyze the LIT allows teachers to track progress in their instructional responses, e.g., use of academic language by students versus by the teacher.

Developing teachers’ sense of efficacy or confidence that they can meet student learning needs is a significant challenge for culture change. Convention has it that student failure often is rooted in difficult family conditions, personality traits like “laziness,” and personal troubles that derail their academic progress. Such accounts of poor stu-

dent outcomes are common among teachers in teams that lack a solid design and support for inquiry. Related are assumptions that student learning is linear and that a student performing well below grade level can never catch up. According to teacher reports, the greatest resource for changing their beliefs was the students themselves. Once a team had succeeded in improving target students’ performance on a specific skill or academic practice, such as writing a coherent paragraph, they saw that the students could learn to be successful and that their instruction had made a difference. Team presentations of results to colleagues helped to discredit attributions of student failure to factors outside school and move the culture toward a sense of collective efficacy and empowerment to make a difference.

Promising school outcomes for NYC’s inquiry model encourage the system’s continued investment in this capacity-building strategy. Evidence of particular challenges schools face in implementing the model frame an agenda for network and school leadership development.

Challenges and Issues for System Leadership

Leaders at all system levels—the DOE, cluster, network, and school—must navigate challenges to staying the course with inquiry-based reform:

- The pull of competing paradigms
- Balancing top-down guidance with bottom-up initiative
- Learning and change in an accountability environment
- Diversity in school readiness
- Accountability demands from state and federal authorities

DOE leaders have been thoughtful and strategic in navigating these challenges over the first four years of the inquiry initiative. Their experience offers important lessons for other districts. First, the inquiry approach to instructional improvement has been given top priority as an engine for school change. School administrators and teachers are empowered and held accountable for making instructional decisions based on evidence of student learning—without distractions of top-down curriculum mandates or professional development initiatives. Second, NYC provides teacher teams with a tested model for inquiry and supports the practice with a rich data system and websites for sharing effective practice. Third, system leaders are attuned to the challenge of balancing guidance and accountability for inquiry with care and respect for professional judgment, innovation, and leadership at the school level; and they have developed mechanism for learning from school practice. Fourth, the DOE has reallocated much of its central office staff and professional development resources toward network leaders and facilitators skilled in developing an inquiry culture.

In sum, the system's significant ambition to develop school capacity to use data for improvement is being implemented with coherent, strategic efforts to support this change. Nonetheless, NYC system leaders grapple with challenges entailed in sustaining momentum on this system reform agenda.

Pull of Competing Models and Paradigms for Instructional Improvement

System leaders encounter strong pressures to pursue well-established alternative approaches to instructional improvement. Prominent among them are curriculum mandates, professional development to promote particular pedagogical content knowledge and skills, or programs focused on whole child development. Each has historical roots and proponents in NYC, as well as in other districts pursuing inquiry-based reform.

Acknowledging that inquiry-based reform competes with other improvement strategies is important to building buy-in at all system levels. Proponents of the inquiry initiative need to communicate often why this reform strategy has priority over popular alternatives. In what ways is this approach coherent with the broader NYC reform strategy of school empowerment and accountability? What is the evidence that it works to build school capacity for continuous improvement? What long-term vision warrants significant investment in the inquiry approach to school improvement?

Investing in collaborative inquiry as the leading school reform strategy means backing off from other approaches in the short run, but not in the long run. Indeed, evidence from mature SAM schools suggests that teacher readiness for professional development for instruction is generated through inquiry and that teacher learning under these conditions is more likely to be translated into practice.

Balance Between Top-Down Guidance and Initiative at the Bottom

The success of a system inquiry initiative depends fundamentally on the commitment of school leaders and their ownership of the initiative. Leading school reform from the top of the system runs the risk of engendering compliance responses and ritual conformity to "requirements." This possibility is especially troublesome for an initiative aiming to change school culture. How school administrators and teachers perceive and understand the policy intent and how well it fits with their reform preferences and leadership make all the difference. Also important is whether or not school and teacher leaders have adequate learning opportunities and resources to lead inquiry-based change in the school.

NYC administrators and staff are pursuing a "professional" over a "bureaucratic" approach to system change.¹⁶ They avoid mandates, requirements, and accountability mechanisms that are likely to engender a compliance stance or resistance among school staff. Instead they convey in communications and meetings their respect for, and interest in, educators' views on many facets of the broader reform effort. Several

times since rolling out the inquiry initiative in 2007–2008 DOE leaders have organized dialogue sessions with school leaders about the approach and resources.

The idea of *reciprocity of accountability* guided the Inquiry initiative's early roll-out. When the DOE began to hold schools accountable, they at the same time asked school leaders to hold the DOE accountable for providing support for their efforts. In particular, when the Quality Review was introduced to evaluate how well schools were using data in making decisions about instruction and system change, the DOE emphasized that the tool was giving the school valuable information about what was and wasn't working well and why. In turn, the DOE has organized numerous meetings with principals over the years to get their feedback and input on the Inquiry initiative.

System leaders face a dilemma. How do they keep the ballast of a clear reform model while ensuring ownership at the school level? How do they weigh the importance of keeping a strong reform model to leverage change against the risk of losing school commitment by over-specifying reform guidelines? If system reform leadership is mainly a problem of teaching and learning, then leaders have to engage the schools where they are. Yet a constructivist approach to policy formation that starts out by accommodating a wide diversity of school readiness could fail to define a strong enough curriculum to leverage and support change. This would increase school ownership but weaken the scaffold for inquiry practice. The challenge for sustaining an inquiry initiative is being clear on its first principles and supporting consistent adaptations.

Accountability Threats to Learning and Change

The DOE's accountability system calls for results, while its Inquiry initiative calls for learning (see chapter 4 for discussion of tensions). For the purpose of this analysis, it is important to note that designs for collaborative inquiry pull teachers away from the comfort of their closed classroom doors and instructional routines and ask them to take the risks of working with colleagues and committing to turning around struggling students. A natural response to performance pressure is to stick with the tried-and-true—to work harder and longer rather than better and smarter with struggling students. The demands of collaborative inquiry may seem entirely too risky to some teachers or they just may feel that they can more efficiently meet demand for improved student outcomes on their own.

The DOE's collaborative inquiry tools are designed to support teacher learning and change in this accountability context. They provide exemplars of team inquiry, evidence of its success, and all sorts of guidelines for team practice. By providing teams with control over access to the information they post about their own practice, the website provides a safe place for sharing across schools and conveys that the site is in no way designed for teacher evaluation purposes.

NYC principals may feel especially exposed and vulnerable to the risk of failing with a new reform model, particularly if their school's professional culture is weak

in social trust. Principals may need incentives and assurance of safety in taking this work on seriously.

The QR is designed partly for this purpose, in that high ratings on leading indicators of a school inquiry culture offer some buffer against sanctions for low Progress Report ratings on student outcome indicators. In such instances, the principal and school are rewarded for moving in the right direction.

Diversity in School Readiness

Readiness to implement collaborative inquiry in teacher teams, and the level of support required, varies widely across schools. Quality Reviews and network leader ratings on the DOE's inquiry capacity measures help to define school readiness and progress. However, it is not clear that school differences on these measures as currently constructed capture the developmental stages or trajectories of school change. Nor is it clear how the information should be used to strengthen inquiry in a given school context.

Research offers little guidance on these issues. Although we have documented the broad arc of change in the way a team works together and how deeply and well they diagnose and address student learning needs, we know much less about what it takes to bring about qualitative shifts toward successful collaborative inquiry. What specific facilitator moves or supports from a principal or tools can move a team beyond the plateaus and roadblocks they encounter? Through what stages and strategies does a school reach a tipping point where inquiry norms overtake resistance to change? Answers to these questions would provide a knowledge base to help focus school leaders' strategic approaches to facilitating inquiry-based reform for schools at different developmental stages and to help define a central office role in fostering school culture change.

Accountability Demands from State and Federal Authorities

New York City's inquiry approach to building capacity for school improvement is in some ways out of sync with the state and federal reform model, and schools experience tensions in the dual accountability systems. The DOE's investment in developing teachers' capacity to use data to improve instruction means that QR ratings (leading indicators) can compensate for negative No Child Left Behind accountability indicators (trailing indicators). Further, a school scoring high on the progress report can be in *schools under registration review* (SURR) status according to New York criteria for subgroup gains, as was the case in one high school we followed. This disjuncture may or may not pull a school away from collaborative inquiry, depending on whether the principal sees it as an effective strategy to improve student test scores, but it certainly narrows attention to test results.

At the national level, rollout of Common Core State Standards (CCSS) may pose a threat to NYC's collaborative inquiry initiative, at least in terms of competing

demands for teachers' focus and learning. In theory, national standards could enrich the assessments and student performance data that teacher teams use to identify and address student learning needs. However, the CCSS initiative will require that teachers learn and understand the new set of standards, design or use curricula that promote the new learning outcomes, and learn how to assess student learning under them. If not managed well, this agenda has the potential to shift attention from student learning back onto adult learning and derail a district's inquiry initiative. NYC has designed a rollout of the Common Core standards and training that articulates with teachers' work in grade-level inquiry teams, calling on some teacher groups in each school to take the lead. It aims to marry grade-level inquiry that empowers teachers' instructional decisions with some teacher teams' development of new standards-based assessments. Both will drive inquiry-based reform in NYC schools.

System leaders committed to sustaining and deepening collaborative inquiry in NYC schools are challenged on several fronts. They navigate internal reform politics that threaten to dilute the effort, grapple with technical and organizational complexities of supporting school change, and respond to multiple accountability and reform demands from broader policy contexts. Whether the system stays the course, or is pulled away from its approach to developing schools' capacity for continuous improvement, depends on how the DOE and other system leaders contend with these considerable challenges.

16. See James P. Spillane, Brian J. Reiser, and Louis M. Gomez, "Policy Implementation and Cognition: The Role of Human, Social, and Distributed Cognition in Framing Policy Implementation," in *Confronting Complexity: Defining the Field of Education Policy Implementation*, ed. Meredith I. Honig, (Albany, NY: The State University of New York Press, 2006).
17. *Ibid.*
18. For example, see Milbrey W. McLaughlin and Joan E. Talbert, *Building School-based Teacher Learning Communities: Professional Strategies to Improve Student Achievement* (New York: Teachers College Press, 2006).
19. This is not to imply that the central office was monitoring this relationship prior to the Bloomberg/Klein administration—only that such monitoring could be helpful for improvement and for equity, particularly in an empowerment-based system.
20. These concerns were exacerbated by the use of the percentage of students at or above proficiency as the measure of progress, as this indicator is highly sensitive to the placement of the proficiency cutoff. This indicator can either overreport progress (if a large number of students move just above the proficiency cutoff) or underreport progress (if scores of students below the cutoff increase but not sufficiently to move them over the proficiency cutoff). Analysis of scale score trends indicate that achievement gaps have indeed been reduced, though substantial inequities in outcomes among student groups remain.
21. In addition to emphasizing access to common standards, the DOE has launched a series of targeted efforts to address the specific needs of these groups. For students with disabilities, the new division has set the expectation that all schools will make flexible use of instructional placements and organizational supports to best meet all students' needs. For ELLs, the DOE has also committed to holding schools and principals accountable for ELL student outcomes while providing targeted support to schools at the network level. Other initiatives include the hiring of parent coordinators in all schools who communicate with parents in their native languages, providing assessments in native languages, and requiring schools to submit reports on services provided to ELLs. According to a senior DOE administrator, "In a big-picture way, we have never paid as much attention to ELLs in the system as we have under Children First." While such initiatives signal a commitment to improving outcomes for these two under-served groups, the challenge will be to ensure that services are integrated into the work of the other divisions in the DOE and into the networks and schools themselves.

Chapter 6

Research described in this chapter was conducted by Stanford University's Center for Research on the Context of Teaching (CRC) through a grant to New Visions for Public Schools from the Carnegie Corporation of New York. CRC colleagues involved in this study are Milbrey McLaughlin, Lambrina Mileva, Pai-rou Chen, M. Ken Cor, and John Schoener. We are grateful to Liz Gewirtzman and Nell Scharff of the Baruch College School of Public Affairs for opening their SAM program practice to our scrutiny and collaborating with us on making sense of the data. We thank New Visions' Robert Hughes, Ron Chaluisan, and Beverly Donohue for actively supporting our research on inquiry in their PSO schools. We are especially grateful to the teachers and principals in approximately eighty NYC schools who have filled out our surveys and talked with us over the past five years, and to the NYC Department of Education leaders we interviewed for this chapter. Several reviewers' comments on an earlier draft helped to get the story right. Thanks to Robert Hughes, Libi Gil, Louis Gomez, and Shael Polakow-Suransky for their helpful feedback on an earlier draft and special thanks to Irma Zardoya and Liz Gewirtzman for their input on multiple drafts. Opinions expressed here are those of the author and do not necessarily reflect the views of funders or reviewers.

1. Stanford University's Center for Research on the Context of Teaching (CRC) conducted the research under the auspices of New Visions for Public Schools. Schools participating in the research include those involved in an inquiry-focused administrator credentialing program since 2005–2006 and all schools in the New Visions for Public Schools Partnership Support Organization (PSO) since it was established in 2007–2008 (approximately seventy-five schools). For information on the design and results of this evaluation research, see Joan E. Talbert and Nell Scharff, "Leading School Improvement with Data: A Theory of Action to Extend the Sphere of Student Success" (paper presented at the Annual Meeting of the Education Research Association, New York City, 2008); Joan E. Talbert, Lambrina Mileva, Milbrey McLaughlin, John Schoener, M. Ken Cor, Pai-rou Chen, and Wendy Lin, *Leadership Development and School Reform Through the Scaffolded Apprenticeship Model (SAM)* (Stanford, CA: Center for Research on the Context of Teaching, 2009); Joan E. Talbert, Lambrina Mileva, Pai-rou Chen, M. Ken Cor, and Milbrey McLaughlin, *Developing School Capacity for Inquiry-based Improvement: Progress, Challenges, and Resources* (Stanford, CA: Center for Research on the Context of Teaching, 2010).

2. Our prior and concurrent research on district system initiatives to promote inquiry-based school reform includes a five-year study (2001–2006) of the Bay Area School Reform Collaborative (BASRC) district reform initiative and ongoing study of Sanger USD in the California Central Valley (2008–). For a distillation of findings from prior research, see Milbrey W. McLaughlin and Joan E. Talbert, *Building School-based Teacher Learning Communities: Professional Strategies to Improve Student Achievement* (New York: Teachers College Press, 2006).
3. The SAM certification program was launched by Liz Gewirtzman of Baruch School of Public Affairs (SPA) in partnership with Ron Chaluisan and Robert Hughes of New Visions for Public Schools. Others involved in the SAM planning process during 2004–05 included Shael Polakow-Suransky, former NYC principal and chief academic officer as of January 2011. He worked with Jim Liebman, Eric Nadelstern, and Alisa Berger in bringing the inquiry model into the DOE's Children First Intensive (CFI). SAM's inquiry model was developed during the pilot year and refined through work with fourteen high school teams during SAM II (2006–2008). SAM leaders contributed lessons gleaned from implementing the administrator credentialing program, as well as tools and professional development, to the DOE's evolving inquiry initiative.
4. For elaboration on this principle and illustrations of how it works in practice, see Helen A Scharff, Deirdre A DeAngelis, and Joan E Talbert, "Starting Small for Big School Improvement: Focusing on Small Changes That Meet the Needs of Struggling Students Can Lead to Changes in Schoolwide Practices," *Principal Leadership* 10, no. 8 (2010): 58.
5. For a map of resources currently available to teacher teams see the DOE's website to support Collaborative Inquiry: New York City Department of Education, Children First Intensive WebSite, <http://cfi.sharepointsite.net>.
6. As part of this restructuring, the DOE created networks of approximately twenty-five schools each within the SSOs, each guided by a network leader. See introduction and O'Day and Bitter (chapter 5 in this volume).
7. See Marian Robinson, Patricia Kannapel, Joan Gujarati, Hakim Williams, Andrea Oettenger, *A Formative Study of the Implementation of the Inquiry Team Process in New York City Public Schools: 2007–2008 Findings* (New York City: Teachers College, Consortium for Policy Research in Education, 2008).
8. During the second year (2008–2009), the DOE asked schools to establish two or more school teams and communicated the two-year goal of involving 90 percent of teachers in inquiry teams.
9. For a practice-based account of what it takes for a principal to create a school culture of collaborative inquiry, see Nancy Mohr and Allen Dichter, *Stages of Team Development: Lessons from the Struggles of Site-Based Management*, 2001, http://www.nsrhfharmony.org/auth_mohr_dichter_stages.pdf. For evidence of a peer facilitator's role in developing a team culture of collective efficacy, see Ronald Gallimore, Bradley A. Ermeling, William M. Saunders, and Claude Goldenburg, "Moving the Learning of Teaching Closer to Practice: Teacher Education Implications of School-based Inquiry Teams," *Elementary School Journal* 109, no. 5 (2009): 537–553.
10. Questions of validity of our research on New Visions PSO schools for documenting NYC's inquiry initiative need to be addressed. First, we do not use these data to estimate the distribution of NYC schools on inquiry implementation or its outcomes. Second, we examined QR ratings for all NYC schools identified by their SSO to assess how well our data from New Visions PSO schools capture the range of school experiences in implementing the inquiry model. New Visions PSO overall school ratings fell in the middle of the distribution (graphic summary available on request). Thus there is no reason to question the reliability of statistical estimates based on our data or the use of case study data from New Visions PSO schools to investigate teachers' experiences in implementing the model.
11. For further analysis of the challenge of developing collaborative practice in high schools, see Milbrey W. McLaughlin and Joan E. Talbert, "Building Professional Learning Communities in High Schools: Challenges and Promising Practices," in *Professional Learning Communities: Divergence, Detail, and Difficulties*, ed. Louise Stoll and Karen Seashore Louis (Maidenhead, UK: Open University Press/McGraw-Hill, 2007).
12. This survey scale combines teacher responses to two items measured on 5-point Likert scales: we use a variety of assessment strategies to measure student progress, and we use assessment data to evaluate our curriculum and instructional practices. The scale's alpha coefficient is .82.
13. A forthcoming monograph documents the developmental trajectories and facilitator moves that support these cultural changes through inquiry: Nell Scharff and Joan E. Talbert, *What It Takes to Develop Evidence-based Practice in Schools: A Developmental Perspective*, in process.
14. Coefficients are .32 ($p = .5$) and $-.47$ ($p = .1$), respectively. School N for the analysis is 38. Results of both student outcome analyses are reported in more detail in Talbert et al., *Leadership Development and School Reform through the Scaffolded Apprenticeship Model (SAM)*.

15. The "principal support of inquiry team" scale (Alpha = .88) combines Likert scale responses to three survey items: Principal: establishes conditions for trust and open communication; actively supports our risk-taking; uses authority to push our learning in the service of target students and targeted learning goals. The "team functioning" scale (Alpha = .90) combines three items: Our inquiry team members: establish clear and unambiguous measures for assessing our success; stay focused on results in the face of distractions and competing priorities; willingly make sacrifices for the good of the team and the achievement of our goals. Principal support predicts 2009 IT scores on the team functioning scale after 2008 scores are controlled (coefficients are .78 ($p = .01$) and .35 ($p = .05$), respectively. Translated: an IT's functioning improved most in schools where the principal promoted the work.

16. Joan E. Talbert, "Professional Learning Communities at the Crossroads: How Systems Hinder or Engender Change," in *International Handbook of Educational Change*, vol. 2, ed. Michael Fullan, Andy Hargreaves, and Ann Lieberman (Dordrecht, The Netherlands: Springer, 2010).

Chapter 7

We appreciate helpful comments on an earlier draft from Susan Moore Johnson, Vicki Bernstein, and participants at the New York City Retrospective working conference. We are grateful to the New York City Department of Education and the New York State Education Department for the data employed in this paper. Loeb and Wyckoff also gratefully acknowledge the support of the National Center for the Analysis of Longitudinal Data in Education Research (CALDER). CALDER is supported by IES Grant R305A060018 to the Urban Institute. The views expressed in the paper are solely those of the authors and may not reflect those of the funders. Any errors are attributable to the authors.

Some sections of this chapter are taken from Margaret Goertz and Stephanie Levin, "Strategic Management of Human Capital in New York City," 2008, <http://www.smhc-cpre.org/download/36/>.

1. *Recruitment efforts*: Donald Boyd, Hamilton Lankford, Susanna Loeb, Jonah Rockoff, and James Wyckoff, "The Narrowing Gap in New York City Teacher Qualifications and Its Implications for Student Achievement in High Poverty Schools," *Journal of Policy Analysis and Management* 27, no. 4 (2008): 793–818. *working conditions*: Susan Moore Johnson, Jill Harrison Berg, and Morgaen L. Donaldson, *Who Stays in Teaching and Why* (Cambridge, MA: The Project on the Next Generation of Teachers, Harvard Graduate School of Education, 2005).

2. Donald Boyd, Pamela Grossman, Hamilton Lankford, Susanna Loeb, and James Wyckoff, "How Changes in Entry Requirements Alter the Teacher Workforce and Affect Student Achievement," *Education Finance and Policy* 1, no. 2 (2006): 176–216.

3. Based on calculations by authors using data from NYCDOE, NYSED, and College Board.

4. Donald Boyd, Hamilton Lankford, Susanna Loeb, and James Wyckoff, "Explaining the Short Careers of High-Achieving Teachers in Schools with Low-Performing Students," *American Economic Review Proceedings* 95, no. 2 (2005): 166–171.

5. Hamilton Lankford, Susanna Loeb, and James Wyckoff, "Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis," *Educational Evaluation and Policy Analysis* 24, no. 1 (2002): 38–62.

6. *Initial search for schools*: Donald Boyd, Hamilton Lankford, Susanna Loeb, Matthew Ronfeldt, and James Wyckoff, "The Role of Teacher Quality in Retention and Hiring: Using Applications-to-Transfer to Uncover Preferences of Teachers and Schools," *Journal of Policy Analysis and Management* (forthcoming); and Donald Boyd, Hamilton Lankford, Susanna Loeb, and James Wyckoff, "Analyzing the Determinants of the Matching of Public School Teachers to Jobs: Estimating Compensating Differentials in Imperfect Labor Markets," NBER Working Paper 9878 (2006); *job retention*: Boyd et al., "The Role of Teacher Quality in Retention and Hiring"; and Eric A. Hanushek, John F. Kain, and Steven G. Rivkin "Why Public Schools Lose Teachers," *Journal of Human Resources* 39, no. 2 (2004): 326–354.

7. Donald Boyd, Pamela Grossman, Marsha Ing, Hamilton Lankford, Susanna Loeb, and James Wyckoff, "The Influence of School Administrators on Teacher Retention Decisions," *American Education Research Journal* (forthcoming); Donald Boyd, Pamela Grossman, Hamilton Lankford, Susanna Loeb, Jeannie Myung, and James Wyckoff "Mentoring and the Matthew Effect: The Implementation of a District-Wide Mentoring Program and its Impact on Retention" (working paper, 2009); Boyd et al., "Explaining the Short Careers of High-Achieving Teachers in Schools with Low-Performing Students"; Hanushek, Kain, and Rivkin, "Why Public Schools Lose Teachers,"; and Johnson, Berg, and Donaldson, *Who Stays in Teaching and Why*.

8. Sean P. Corcoran, Amy Ellen Schwartz, and Meryle Weinstein, *The New York City Aspiring Principals Program: A School-Level Evaluation* (report from Institute for Education and Social Policy, New York University, 2009), <http://steinhardt.nyu.edu/scmsAdmin/uploads/003/852/APP.pdf><http://steinhardt.nyu.edu/scmsAdmin/uploads/003/852/APP.pdf>.

9. The New Teacher Project, "Unintended Consequences" (2005), <http://www.tntp.org/files/Unintended-Consequences.pdf>.

10. Based on calculations by authors from data provided by NYCDOE.

11. See the Agreement Between the Board of Education of the City School District of the City of New York and the United Federation of Teachers, 2000-2003 Article 18 Transfer and Staffing, 120–122.

12. *Education Week*, January 6, 2005, 86.

13. Elissa Gootman, "A New Effort to Remove Bad Teachers," *New York Times*, November 15, 2007, http://www.nytimes.com/2007/11/15/education/15teacher.html?_r=1http://www.nytimes.com/2007/11/15/education/15teacher.html?_r=1

14. Stephen Brill, "The Rubber Room: The Battle over New York City's Worst Teachers," *New Yorker*, August 31, 2009, http://www.newyorker.com/reporting/2009/08/31/090831fa_fact_brill.

15. Julia Koppich and Connie Showalter, *Strategic Management of Human Capital: A Cross-Case Analysis of Five Districts* (Madison, WI: University of Wisconsin-Madison, Consortium for Policy Research in Education, 2008), <http://www.smhc-cpre.org>.

16. Most of this increase occurred during the period 2000–2002. Between 2002 and 2008 starting salaries increased by 17 percent. After adjusting for inflation this translates to a real increase in salary of less than 0.5 percent.

17. For a complete list of the eligibility requirements see <http://schools.nyc.gov/NR/rdonlyres/9720279E-BD06-464A-88F3-ADF67CABB29E/0/LeadTeacherPosting20102011A.pdf>.

18. Matthew G. Springer and Marcus A. Winters, "New York City's School-Wide Bonus Pay Program: Early Evidence from a Randomized Trial" (working paper, National Center on Performance Incentives, Vanderbilt University, 2009), http://www.performanceincentives.org/data/files/news/PapersNews/Springer__Winters_200902.pdf.

19. Donald Boyd et al., "The Narrowing Gap in New York City Teacher Qualifications."

20. Charles Clotfelter, Elizabeth Glennie, Helen Ladd, and Jacob Vigdor, "Would Higher Salaries Keep Teachers in High-Poverty Schools? Evidence from a Policy Intervention in North Carolina." (NBER Working Paper #12285, 2006); Peter J. Dolton and Wilbert van der Klaaw, "The Turnover of Teachers: A Competing Risks Explanation," *Review of Economics and Statistics*, 81, no. 3 (1999): 543–552; Richard J. Murnane, Judith D. Singer, and John B. Willett, "The Influences of Salaries and 'Opportunity Costs' on Teachers' Career Choices: Evidence from North Carolina," *Harvard Educational Review*, 59, no. 3 (1989): 325–346; Todd R. Stinebrickner, "A Dynamic Model of Teacher Labor Supply," *Journal of Labor Economics*, 19, no. 1 (2001): 196–230; Todd R. Stinebrickner, "Compensation Policies and Teacher Decisions," *International Economic Review*, 42, no. 3 (2001): 751–779.

21. Donald Boyd, Pamela Grossman, Karen Hammerness, Hamilton Lankford, Susanna Loeb, Matthew Ronfeldt, and James Wyckoff, "Recruiting Effective Math Teachers: How Do Math Immersion Teachers Compare? Evidence from New York City" (NBER Working Paper w16017, 2009).

22. Even more dramatic differences exist on other measures of qualifications, such as undergraduate college ranking and the percentage who failed the Liberal Arts and Sciences Test (general teacher certification exam) on their first attempt. See Boyd et al., "The Narrowing Gap in New York City Teacher Qualifications."

23. Boyd et al., "Recruiting Effective Math Teachers"; Boyd et al., "Analyzing the Determinants of the Matching of Public School Teachers to Jobs"; and Thomas J. Kane, Jonah E. Rockoff, and Douglas O. Staiger, "What Does Certification Tell Us About Teacher Effectiveness? Evidence from New York City" *Economics of Education Review* 27 (2008): 615–631.

24. For details see Boyd et al., "How Changes in Entry Requirements Alter the Teacher Workforce and Affect Student Achievement"; and Boyd et al., "Recruiting Effective Math Teachers."

25. Boyd et al., "Recruiting Effective Math Teachers."

26. Boyd et al., "The Narrowing Gap in New York City Teacher Qualifications."

27. *Ibid.*

28. *Ibid.*

29. Jonah E. Rockoff, Brian A. Jacob, Thomas J. Kane, and Douglas O. Staiger, "Can You Recognize an Effective Teacher When You Recruit One?" (NBER Working Paper 14485, 2008).

30. This section is based on conversations with Vicki Bernstein, executive director of teacher recruitment and quality, NYCDOE.

31. Jason A. Grissom, "But Do They Stay? Addressing Issues of Teacher Retention through Alternative Certification," in *Alternative Routes to Teaching: Mapping the New Landscape of Teacher Education*, eds. Pamela