Center for Research on the Context of Teaching
Stanford University

CRC

Inquiry-based School Reform:
Lessons from SAM in NYC

Joan E. Talbert
with
M. Ken Cor
Pai-rou Chen
Lambrina Mileva Kless
Milbrey McLaughlin

Final Report

January 2012

Table of Contents

Introduction

Policymakers, district and school administrators, educators, and researchers generally agree that data-based decision making, teacher collaboration, and distributed leadership are important conditions for improving school success. Yet, rarely have systems been able to establish these conditions across many schools. The rhetoric associated with these improvement strategies travels well, however, "the underlying conceptualization and thinking do not" (Fullan, 2005, p. 10). Moreover, studies of schools that effectively use collaborative inquiry to improve student achievement fall short of explaining how they developed this capacity. Challenges of change and ways in which internal and external leaders support progress toward inquiry-based school reform are little understood.

This four-year evaluation of inquiry-based school reform in New Visions for Public Schools' network addresses this knowledge gap.[1] Since 2007-08, Stanford researchers have documented schools' experiences implementing the Scaffolded Apprenticeship Model (SAM) for inquiry-based school improvement. SAM promotes teams' use of a particular inquiry model to address student skill gaps, fosters distributed inquiry leadership, and includes an external facilitator to support teachers' development of collaborative inquiry and leadership skills. Pioneered by New Visions and Baruch College School of Public Affairs as an administrator credentialing program, SAM became a model for New York City's *Collaborative Inquiry* initiative under Children First. Over the past four years, all NYC schools were expected to establish teams of teachers to collaborate in using data on student performance to improve their success. Some schools sent teacher teams to participate in the SAM credentialing program, which joins leadership development through a university-based, degree-granting program with inquiry-based school reform. The certification program provides intensive scaffolding and facilitation of teams' implementation of the inquiry model and inquiry leadership with school colleagues.

The evaluation uses SAM's theory of change as its framework and analyzes school implementation patterns and outcomes. Central research questions are:
- To what extent do schools that implement collaborative inquiry improve student success rates? Does the credentialing program make a difference in outcomes?
- How does collaborative inquiry change teaching and school culture? What are the challenges of implementing SAM and what resources make a difference?
- What are the challenges of facilitating SAM implementation and what makes a difference for effectiveness?

The issue of whether and how the SAM credentialing program helps schools implement inquiry-based reform has significant resource implications. If schools involved in the program do better at improving student outcomes, after prior student achievement and demographics are taken into account, then the question arises of how the program's results can be achieved at affordable costs without sustained foundation support.

---

[1] The evaluation also provided formative feedback to New Visions and the SAM credentialing program at Baruch College School of Public Policy through annual reports and briefings.

New Visions' network provides an ideal setting in which to investigate schools' implementation of SAM's theory of change toward inquiry-based reform and the role of the certification program in supporting change. The network includes a substantial number of schools that have been implementing the SAM model with support from the credentialing program over the past four years, as well as a large number of schools that have not participated in the program. With funding from the Carnegie Corporation of New York, New Visions administered tuition and school support payments for five successive cohorts of teacher teams to participate in the certification program.[2] Since 2008, 16 of approximately 70 high schools in New Visions' network had at least one teacher team that participated in the program (hereafter "SAM-cert" schools). The remaining (non-cert) schools were implementing NYC's collaborative inquiry model with support from New Visions' facilitators and data team, as well as resources developed by the DOE.[3]

After describing SAM's theory of change and our evaluation design and data, we report results of statistical analyses of student outcomes to address the question of SAM-cert effects. We then focus on implementation – the ways in which collaborative inquiry influences teaching and school culture, challenges school teams encounter in developing this practice, how principals and facilitators make a difference for school progress, and challenges of facilitating change toward inquiry-based school reform. Given evidence of significant SAM-cert effects in our analysis of student outcomes, we consider particular resources for change afforded by the program in these analyses of implementation. The last section highlights key findings from our four year evaluation of the SAM model and what they suggest about the promise and prospects for sustaining and deepening school improvement through inquiry-based reform in New Visions schools and NYC more broadly.

---

[2] SAM III was launched in January 2008 and included successive cohorts of teacher teams from 16 schools in New Visions' Network, including four schools that had been involved in SAM II; teacher teams from several schools outside New Visions network also were part of SAM III. Nell Scharff Panero of Baruch College trained six SAM III instructors to work with these schools. The final cohort of SAM III teams will complete the program in January 2012.

[3] See Talbert (2011) for discussion and details of NYC's design and supports for school implementation of SAM's model for collaborative inquiry.

SAM's Theory of School Change and Design

SAM posits that collaborative inquiry – in which a team of teachers uses data on student performance to identify specific learning gaps and target interventions to address them – will continuously improve student success. Its stance on the problem of change extends beyond many inquiry-based reform models' emphasis on technical challenges of developing useful student data systems and teachers' skills in using assessment data – to include school organization and culture. SAM assumes that every school has a "sphere of success," or a group of students with whom the school is currently successful, and the challenge of change is tackling assumptions and instructional practices and decision systems that keep students outside this sphere. SAM engages a team of teachers in systematically using evidence of struggling students' skill gaps to both design instructional responses and re-design systems that inhibit their skill development. This approach to school improvement seems fairly straightforward. However, it challenges current practices and thinking about quality education, especially a common notion that high content performance standards means that some students will fail.

SAM's model for inquiry-based school reform is designed strategically to leverage particular shifts in teachers' thinking about why students struggle, whether and how teachers can accelerate their learning, and how schools might reduce barriers to their progress. The credentialing program additionally prepares individuals to lead school change through coaching colleagues toward an inquiry stance and facilitating school inquiry teams, as well as designing instructional decision systems to better support student success. At its core is the design for teacher team inquiry that was adopted and refined by NYC's collaborative inquiry initiative.

SAM's design features *teams, tasks, and tools* to develop inquiry practices and move a school culture toward conditions for continuous improvement. Each is designed to challenge habits of mind that constrain some students' success in a school and to develop skills in using inquiry to improve their achievement. The design also calls for a facilitator to guide the work, keep the team focused on student learning, and nurture shifts in beliefs toward collective efficacy or the belief that collaborative inquiry can bring about improved student achievement.

Inquiry Team. In its original design for developing school leadership capable of addressing school failures, SAM called for a school "Inquiry Team" comprised of teachers who represented a broad array of units (e.g., subject departments or grade levels). This team composition was a vehicle for distributing leadership in the school. At the same time, it created a community of practice around the work of improving student learning through data-based inquiry. Under NYC's scaled-up version of SAM, all teachers in a school are expected to form grade-level and/or content area "collaborative inquiry" teams, and the credentialing program has welcomed successive cohorts of SAM-cert school team members with less stringent composition criteria.

SAM-cert and non-cert teams have different levels of support for collaborative inquiry. The program curriculum scaffolds and structures inquiry practice and prompts teams to think systemically about the problem of school change. Program participants learn how to work as a team that uses evidence of student skill gaps to design instructional interventions and system

changes and how to lead culture change with colleagues through modeling a learning stance and sharing evidence of effective interventions.

Tasks to guide inquiry practice. As a model for collaborative inquiry, SAM specifies stages and tasks for a team's collaborative inquiry. Teams are prompted to:
1. Conduct inquiry to move specific students outside the school's sphere of success and change systems that constrain those students;
2. Conduct inquiry to move colleagues toward inquiry practice;
3. Establish systems to ensure sustainability.

These tasks anchor the SAM-cert curriculum and assignments. Team assignments and selected readings support connections between practical problems and conceptual frames and guidelines from relevant literatures within each practice-based module. For example, the task of focusing closely on specific skill gaps of a small group of struggling students prompts team members to shift their thinking from teaching to learning and sets the stage for their purposive reading of formative assessment literature and, ultimately, of literature on leading school culture change.

A core SAM principle for school change is "getting small" in order to go big with inquiry-based improvement (Scharff, DeAngelis, Talbert, 2010). Teams are prompted to start with a small number of students, focus on a specific skill gap – such as reading comprehension – and move the students on a particular learning target (LT) relevant to that gap, such as topic recognition or using context cues.[4] This not only makes the team's work manageable in scope, but also prompts important shifts in team members' perspectives about why students are not successful and how their learning can be accelerated. It shifts their focus toward evidence of struggling students' specific skill gaps and learning needs and away from family circumstances outside the school's control.

Tools to frame and support SAM teams' work. SAM-cert's curriculum requires that teams use multiple tools to identify and address needs for system change in their schools. The tools are designed to support a precise and rigorous focus on *how* current conditions produce current outcomes and *how* the team might change these conditions.

Key SAM-cert tools include:
- Low-inference transcripts (LITs): verbatim scripts of everything that is said in a classroom. SAM participants learn to create and analyze these transcripts through the focused lens of an identified skill gap for target students.
- Readings aligned with SAM curriculum modules and assignments.
- Protocols for coaching and inter-visitations among SAM-cert schools.

Low inference transcripts of target students' classes support a shift in teachers' focus from teaching (as curriculum delivery or instructional practices) to student learning, and open new ways for them to see target students and their potential impact on those students. LITs

---

[4] A learning target is more granular than a skill or subskill. For example, the skill of reading includes the subskill comprehension (as well as phonemic awareness, phonics, fluency, and vocabulary). Comprehension, in turn, includes learning targets such as topic recognition, using context cues, and inferring or drawing conclusions.

provide data that enable a teacher to understand the curriculum as taught and experienced by students – rather than as it exists on a map or in their minds. When teachers analyze transcripts through the specific lens of what target students do not know to see if it is taught to them, they are confronted with the reality that the students typically do not have the opportunity in any of their classes to learn the skills they lack. In this way, LITs support the core task of understanding how current school conditions produce current student outcomes.

Readings are designed to generate new perspectives that have been scaffolded by an experience (rather than vice versa, as is typical of credentialing programs). For example, teams work on developing assessments of a learning target before reading literature on formative versus summative assessment, and they analyze patterns of resistance to inquiry in their school before reading change leadership literature. Carefully selected readings for each SAM module help teachers make particular cognitive shifts essential to the work. For example, during the research phase, they read articles and chapters that address differences between summative and formative assessments (Popham, 2001); during the action phase, they read pieces that support a shift from research to action, such as "Closing the Knowing-Doing Gap" (DuFour, Eaker and DuFour, 2005); and during the sustainable leadership phase they read such pieces as *Leadership on the Line* (Heifetz and Linsky, 2002).

Protocols for facilitator coaching with team members and inter-visitations between schools guide SAM instructors/facilitators' action to support individual, team, and school cultural shifts. Individual leadership coaching helps participants develop strategies for managing what is most difficult for them personally in the work of doing and leading school improvement. The coaching protocol leads the teacher to identify an area of personal challenge, to understand the assumptions and beliefs that create current responses, and to develop a plan for addressing them.

Inter-visitations are designed to support each school team's learning. The host team receives feedback from other SAM participants, and all learn how to provide useful feedback to colleagues. The protocol establishes a structure and norms for a productive visit: the host school defines a problem of change and invites one or more school teams to participate in the problem-solving process. The visiting team collects data and provides an outside perspective on the problem. This practice and protocol prompt a shift toward public learning and help develop leaders' skills in giving honest and actionable feedback.

Facilitators. In SAM-cert and DOE versions of collaborative inquiry, a facilitator is assigned to each school to help inquiry teams' use data to design and implement responses to target students' skill gaps. The facilitator's role is to keep teams on task, provide feedback on their work, and push them to develop the discipline of inquiry. For example, s/he holds the team accountable for assessing student learning outcomes of their intervention and refining it to ensure student mastery of a measurable skill. As noted, SAM-cert facilitators additionally require teams to use data to improve decision-making systems that have led to the identified skill gaps and help individuals develop skills for leading organizational change. The program's weekly seminars ensure regular interaction between school teams and their facilitator and keep the teams focused and on course with their inquiry work. In contrast, facilitators in non-cert schools have a wide range of responsibilities that limit the time and attention they can devote to supporting inquiry

teams' progress. For example, New Visions' Leadership Development Facilitators (LDFs) are essentially point people in the schools, coordinating a wide range of services required by their PSO's contract with the DOE.

SAM facilitators/ instructors work with each school team to translate core principles of the inquiry model into specifics of their school. They guide teams in diagnosing and addressing particular student skill gaps and school conditions that limit target students' success. In order to customize teams' inquiry work, facilitators need to be deeply grounded in SAM's core principles and clear on how its tasks and tools leverage and support their learning. Toward this end, the program has built in a day per week for facilitators' collaborative inquiry. In weekly meetings, facilitators, review teams' work on assignments, design seminar lessons to address skill gaps, and calibrate standards for assessing team and individual work. By all accounts, these practices and the collaborative inquiry they promote have improved the program's quality and facilitators' effectiveness. Further, the facilitators help one another craft strategic responses to whatever challenges they encounter in working with particular schools. They regard their facilitator learning community as crucial to their individual and collective success.

Figure 1 sketches SAM's theory of change for inquiry-based school reform. The diagram shows assumptions about cause-effect relationships between core SAM design features and the intended intermediate organization outcomes and ultimate student outcomes. "SAM Inquiry Model" refers to the particular design for collaborative inquiry – the teams, tasks, and tools just described. School teams involved in the credential program receive more intensive push and support from facilitators/instructors and are held accountable for improving target student achievement and supporting colleagues' inquiry in their school.

Figure 1. SAM's Theory of Change

PARTNERS	SAM MODEL AND DESIGN	IMPLEMENTATION	INTERMEDIATE OUTCOMES	OUTCOMES

Schools

Teacher inquiry teams; common planning time dedicated to collaborative inquiry

Administrators support team inquiry and teacher leadership

Broad leadership of inquiry and data-based decision making for school improvement

New Visions

Support of SAM implementation: Facilitators, data systems and training, Funding agent for SAM-cert

Team(s) implements Inquiry Model

Accelerated gains of struggling students; Expanded sphere of student success

SAM Inquiry Model: Design for data-based inquiry cycles Resources/tools to support the work

School change toward a culture of inquiry to identify and respond to student needs

Baruch College

Credential Program: SAM curriculum and team standards Facilitator/Instructors' intensive training and team support Sustained instructor training

Facilitators guide team learning of inquiry and school leadership practices

Pipeline of certified administrators prepared to lead inquiry-based reform

Evaluation Research Methods, Data, and Analyses

The evaluation was designed to assess the SAM model's assumptions about cause-effect relationships over time, as well as to provide ongoing feedback to SAM and NV leaders on their efforts to support the development of evidence-based school cultures. This final report summarizes evidence concerning:

- o Student outcomes of SAM: expanded sphere of success, in terms of proportions of students on track versus off track for meeting graduation and college readiness standards
- o Intermediate school outcomes: teaching and school culture shifts that the model engenders and implementation challenges
- o Challenges of implementing the facilitator role in SAM and resources that improve effectiveness

Evaluation methods combined quantitative data for all New Visions network schools and longitudinal, field-based case studies of ten high schools. The case study sample was selected to represent critical contrasts of SAM-cert and non-cert schools (5 each) and large and small SAM-cert schools (3 and 2, respectively). Four of the five SAM-cert schools had teams in the SAM II iteration of the program, as well as in at least two of the four SAM III cohorts, affording opportunity to study developmental trajectories of implementation and school culture change.

Data and student outcome measures

This report uses this past year's survey and case-study data to update findings reported in previous reports and new data to assess student outcomes of the SAM program. Data analyzed for this report include:

- student outcomes for the 2012 graduation cohort of all New Visions schools
- annual teacher surveys (2008-2011), including 3-8 teachers in all New Visions schools (members of the school's original Inquiry Team) and all teachers in ten case study schools[5]
- longitudinal case studies in 10 high schools:
  - * semi-annual principal interviews
  - * quarterly focus groups and/or interviews with SAM-cert participants
  - * quarterly focus groups and/or interviews with inquiry leaders in each school, including a person designated as Data Specialist and school inquiry team members
  - * observations of inquiry team meetings in SAM-cert schools
- annual interviews with New Visions facilitators
- annual interviews and observations with SAM facilitators (Baruch College)

Student outcome measures use New Visions' metric for tracking students' progress to graduation and college readiness. The metric includes benchmarks for students' credit

---

[5] Response Ns and rates for each survey year were 2008: 164 (54%); 2009: 296 (64%); 2010: 164 (45%); 2011: 258 (64%).

accumulation and Regents exam scores by semester and indicates where a student needs to be on the graduation and college readiness trajectory over the eight semesters of high school. At any point in a student's high school career s/he can be classified as: a) On Track for College Readiness; b) On Track to Graduate; c) Almost on Track to Graduate; or d) Off Track. This measurement system was designed for high school students and their families, to serve as an early warning system for students who fall behind and to make clear what students need to accomplish to get on track for graduation and college readiness.[6]

The metric is well-suited to our evaluation of SAM because it aims to bring more students into a school's sphere of success. The inquiry model assumes that schools differ in the particular skill gaps and systems that inhibit some students' success, so traditional outcome measures of test scores in particular content areas may or may not capture results. For example, improved Math scores is the goal of some schools, while improved Global Studies scores are the focus of others. In theory, schools use data to define successive skill gaps and interventions to bring more and more students up to grade level standards across content areas. A composite measure of a student's course completion and Regents scores captures the desired outcome – whether or not the school has brought the student into the sphere of success. The school-level outcome is growing proportions of students who are on track for graduation and college readiness.

We focus on student outcomes in 2011 for the 2012 graduation cohort on grounds that they would have had sufficient time in the school to benefit from the progress of inquiry work over the past three years. Their on-track statuses at the end of 2011 would indicate the extent to which the school's inquiry work had a cumulative effect over their time in the school. The 2011 cohort was less desirable mainly because some schools' involvement in the SAM program began in the middle of these students' junior year, affording less opportunity for them to benefit. Also, the scramble for credit recovery during senior year makes their data a less valid measure of learning outcomes of inquiry-based reform.

Measures for the 2012 cohort students are based on their record through the sixth semester of high school. New Visions' benchmarks classify these students as follows:
1) On Track for College Readiness: 33 credits (6 English, 6 math, 6 social studies, 6 sciences, 4 foreign languages, 3.48 PE) *and* 7 Regents passed at 65 and above, including 75 in Math and ELA.
2) On Track to Graduate: 33 credits (same as above) *and* 5 Regents passed at 65
3) Almost on Track to Graduate: 30 credits (4 English, 2 math, 4 social studies, 2 science *and* 2 Regents passed at 65
4) Off track

School outcomes are measured as percent of the 2012 student cohort that was On Track for college readiness or graduation by the end of June 2011.

---

[6] See Fairchild, Gunton, Donohue, Berry, Genn, & Knevals (2011) for more detail on the metric and its intended use to track student progress.

Analyses

       Student outcome analyses focus on the question of whether or not a school's participation in the SAM certification program makes a difference for getting students on track for graduation and college readiness. We test the hypothesis that students in SAM schools do better than expected on the basis of their 8th grade scores on New York State math and ELA tests, their poverty status (eligibility for free and reduced price meals) and race/ethnicity. We use two strategies to assess a SAM effect: 1) regression analysis using a SAM ordinal scale (2-1-0, depending on whether the school had multiple, one, or no teacher teams that participated in the certification program) with controls for student demographics and 8th grade test scores, as well as school size; and 2) matched-sample comparison using a group of non-cert New Visions schools similar in student demographics and 8th grade test scores.[7]

       To analyze *school culture changes* through the SAM model we draw on data for schools that sustained implementation for at least four years. These include four case study schools with successive cohorts of teams that participated in the SAM-cert program – two large restructured high schools and two small schools. Qualitative data reveal patterns in teachers' reports of changes they experienced through collaborative inquiry and changes in school culture that they and administrators reported, as well as strategies that school leaders used to spread and deepen inquiry in the school. Quantitative analysis of teacher survey measures documents six-year trends on teachers' use of student assessment data to inform instruction (see Appendix C for items that make up the "Culture of Assessment Use" scale). Further, for the broader sample of New Visions schools, we estimate effects of this variable and a survey index of school inquiry leadership ("Leadership for Data-based Improvement") on student outcomes. This tests SAM's theory of change linking a school's inquiry culture to expanded student success.

       Our analysis of *implementation challenge and resources* uses both case study data on teams' experiences using inquiry to improve student achievement and quantitative analyses of teacher survey ratings of their principal's and facilitator's support of inquiry across New Visions schools. Qualitative case study data, as well as interviews with facilitators, reveal patterns in teams' struggles with facets of the SAM inquiry model and the particular resources they found to be useful in moving past the hurdles. We test the assumption in SAM's theory of change that a strong facilitator makes a difference for inquiry team functioning and explore the roles that principals play in establishing conditions for collaborative inquiry.

       Evidence of *challenges facilitator face and resources that matter* comes from annual interviews with SAM instructors/facilitators and New Visions facilitators, SAM seminar observations, and school case studies. We identify patterns in facilitator reports of challenges and struggles they encountered in helping teams and schools implement the SAM model, as well as resources that supported their effectiveness. Teacher and administrator reports augment these data for case study schools.

---

[7] Steps to create the matched sample included: 1) Regression analysis for non-cert schools to obtain coefficients for predictors of students' On Track status, 2) Use of regression coefficients to create a school score (referred to as the School Characteristics Index, SCI) for both SAM and non-SAM schools, and 3) Use of the SCIs to match SAM-cert schools with non-cert schools.

Student Outcomes: Does SAM-cert make a Difference?

Our analysis of student outcomes centers on the question of whether the SAM certification program makes a difference for a school's success in getting students on track to graduate and ready for college. We also assess effects on student outcomes of intermediate school outcomes specified in the SAM logic model – inquiry leadership and a culture of assessment use.

New Visions' On Track metric provides a useful index of student achievement that corresponds to SAM's notion of bringing students into the school's "sphere of success." Inquiry-based school reform should result in growing proportions of students completing courses and achieving Regents scores that get them on track for graduation and for college readiness, as well as shrinking proportions seriously off-track for graduation. We analyze students' on-track statuses for the 2012 graduation cohort, since these students would have been in the school for three years during schools' inquiry work and are the cohort most likely to have benefited from school culture shifts.

Transfer schools and schools with fewer than 10 students in the cohort are excluded from the analysis. Based on these exclusion criteria, 45 out of 74 New Visions schools were retained for the main analysis, including all 13 SAM-cert schools and 32 non-cert schools.[8]

SAM program participation effects

First, we estimate effects of a school's involvement in the SAM program on student outcomes at the school level. The SAM Involvement variable captures The SAM variable captures three levels of a school's involvement in the credentialing program: non-cert schools are coded zero (0); schools that participated in SAM since 2009 are coded 2; and schools involved in SAM since 2010 are coded 1.

Two separate regression analyses estimate effects of a school's SAM involvement on: a) Percent Students On Track (proportion of students in the 2012 cohort who are either "On Track for Graduation" or "On Track for College Readiness"); and b) Percent Off Track (excluded from both analyses are the percent students classified as "Almost on Track" for graduation). Each model includes a set of control variables: two cohort-level measures of these students' prior academic performance (percent who scored Below Basic in 8th grade on state standardized tests in reading and in mathematics) and five school-level variables: size (total number of students), poverty (percent students eligible for free or reduced price meals), and race-ethnic composition (percent Asian, Hispanic, and Black students).

On average, schools in which teacher teams had participated in the SAM program for at least a year did better in getting students on track, controlling for students' prior achievement and demographics. Results reported in Table 1 show a statistically significant regression effect of a school's SAM involvement on proportion of students in the 2012 cohort who were "On Track"

---

[8] Two high schools sent teams to the SAM program for the first time in January 2011; they were classified as non-cert for this analysis, since student benefits would not accrue in just half a year.

for graduation or college readiness by June 2011 ($\beta = .39$, $p < .01$). Conversely, a school's SAM involvement predicted a lower percent of students who were seriously "Off Track' ($\beta = -.31$, $p < .05$).

Table 1. Student Outcomes for 2012 Graduation Cohort: Estimating SAM Effects (School N=45)

	Percent On Track[a]	Percent Off Track[b]
SAM Involvement	.39**	-.31*
% Below Basic 8th Grade Reading	.01	-.29*
% Below Basic 8th Grade Math	-.56***	.73***
School Size (Student N)	-.17	.12
Proportion Free and Reduced-price Meals	-.28*	.21
%Asian	-.17	-.15
%Hispanic	-.15	-.04
%Black	-.14	-.08
R^2	.65***	.68***

[a] Percent On Track is the proportion of students in a school classified as 'on track for graduation' or 'on track for college readiness.'
[b] % Off Track is the proportion of students in a school classified as 'off track' (excluding students classified as 'almost on track').
 * $p < .05$, ** $p < .01$, *** $p < .001$

Next we compare student outcome distributions for SAM-cert and matched non-cert schools. This provides a more concrete look at their differential outcomes and interpretation of the regression effect. Figure 2 shows the mean distributions of the 2012 student cohort's statuses on the On Track metric for the SAM-cert schools and matched non-cert schools. These data show that SAM-cert schools far out-perform comparison schools in getting students on track for college readiness and shrinking the proportion of seriously off-track students.

Figure 2. On Track Statuses of 2012 Cohort Students: SAM-cert and Matched non-cert schools
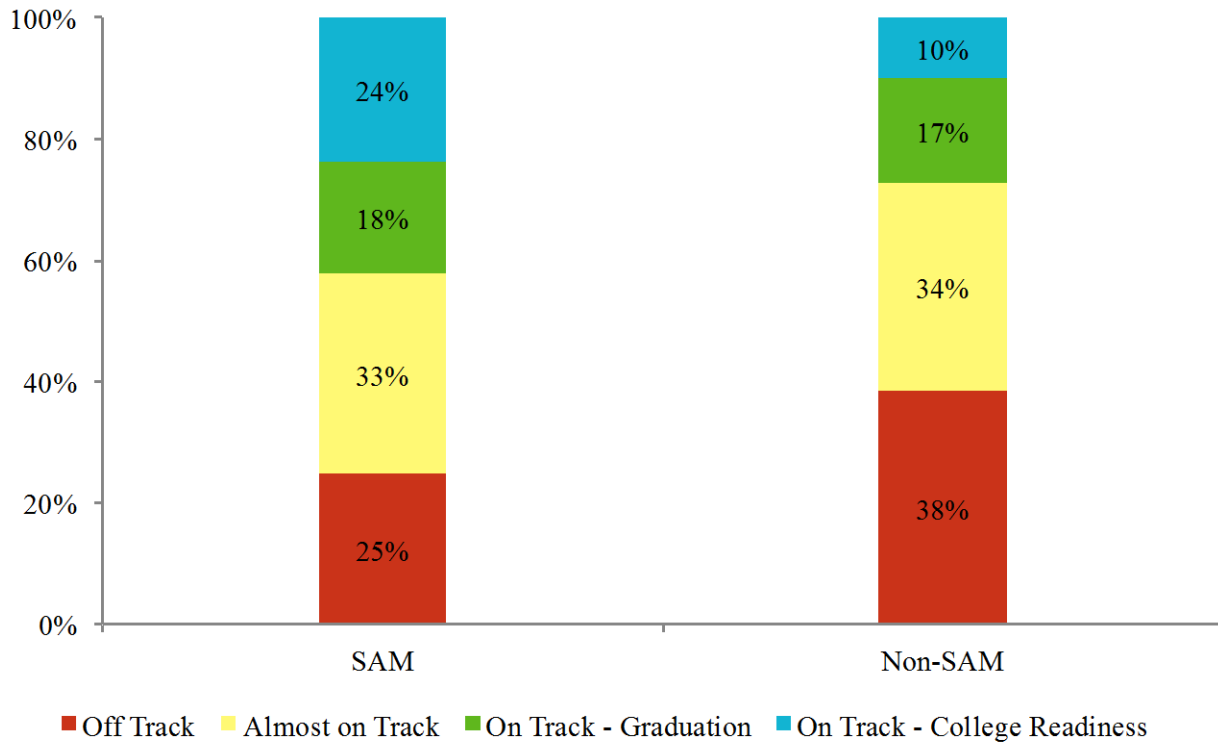
Note: Each group includes 13 schools. Schools in the non-cert group were matched with SAM-cert schools using a School Composition Index (SCI). The SCI is a composite measuring: percent students in the 2012 Cohort who: a) scored Below Basic in Reading in 8th grade, b) scored Below Basic in Math in 8th grade, c) are eligible for Free and Reduced-Price meals, d) are Asian, e) are Hispanic, f) are black, and g) are white. Appendix B shows means on SCI and each variable for the two groups of schools.

The proportion of students on track for colleage readiness after three years in their school was substantially greater in SAM-cert schools than in matched comparson schools – 24 percent versus 10 percent. A substantial difference also shows up in the the proportion of students seriously off track for graduation – 25 percent in SAM-cert schools versus 38 percent in comparison schools. At both ends of the student performance spectrum, schools that invested in the SAM inquiry model and developed teacher leadership through the credentialing program did better that their counterparts.

This comparison under-estimates the SAM-cert effect on student outcomes in the New Visions network. This is because large schools are over-represented among the SAM-cert schools. Tthe number of 2012 cohort students in SAM-cert schools for this analysis is 3288, while the number for matched non-cert schools is 1148. Across New Visions schools included in the regression analyses reported in Table 1, over half of the students were in SAM-cert schools (3288 versus 3237). A student-level regression model (not included here) shows standardized effects of the school's SAM involvement on On Track status that is nearly equivalent to the regression effect of their 8th grade reading score (.15 versus .18) and much greater than their

poverty status (.15 versus -.05). In other words, being in a SAM-cert school increased the likelihood of the student being on track for graduation or college readiness approximately as much as having relatively high 8th grade reading skills.

Last, we test the SAM-cert effect on school outcomes using a hierarchical modeling technique that takes into account the nested character of the data. The model adjusts for any bias that might be attributed to the extraordinary success of two large restructured high schools that invested heavily in the SAM model. The correlation of SAM involvement and school size among New Visions high schools included in this analysis is .54 (see Appendix A).

Table 2. Student Outcomes for 2012 Cohort: Hierarchical Models

	On-Track	Off-Track
Student-level variables:		
8th Grade Reading Score	.18***	-.18***
8th Grade Math Score	.33***	-.30***
School-level variables:		
SAM-cert Involvement	.12*	-.07
School Size (Student N)	-.05	.03
Proportion Eligible for Free/Reduced Price Meals	-.07	.04
%Asian	-.03	-.08
%Hispanic	.00	-.10
%Black	-.02	-.07
R2 (Explained Variance)	.63	.61

Notes. Variable coding is the same as for analysis reported in Table 1. Ns for the analyses are: Student level N= 4919; school level N= 45.
* $p < .05$, ** $p < .01$, *** $p < .001$

Results reported in Table 2 show that a school's involvement in SAM-cert is a significant predictor of its students' On Track status after three years in the school, taking into account both student-level and school-level predictors. However, the hierarchical model for a student's Off-track status yields a non-significant negative effect of their school's SAM involvement.

As noted earlier, controlling for school size in the school-level model yields a conservative estimate ot the SAM program effect, since "star" mature schools are two restructured high schools. Ns here are too small to test for an interaction effect of school size and SAM-cert on student outcomes, but the hypothesis has emerged that SAM might be a particularly effective resource for developing inquiry cultures in large restructuring high schools.

School inquiry culture effects

Regardless of a school's participation in the credentialing program, SAM's theory of inquiry-based reform argues that a school's leadership for data-based inquiry and teachers' use of student assessment data to focus instruction are engines for improving student achievement. To investgate these hypotheses, we used survey measures of these school condtions from our 2011 teacher survey in the New Visions network schools. School outcomes for students' On Track statuses for the 2012 cohort were regressed on survey measures of Culture of Inquiry (COI) and Leadership for Data-based Improvement (LDBI), with controls for student composition and school size. (Appendix C shows teacher survey items that make up these scales and Alpha coefficients).

Table 3 shows results for two regression models for On Track and Off Track student outcomes – one with only the school culture predictors and one that includes the SAM Involvement measure. Model I shows regression effects for the two school culture variables that are not statistically significant. However, the LDBI effect approaches significance for both measures of student outcomes (p < .10). Since SAM participation and LDBI are uncorrelated (r=.19 NS), this effect does not *account for* the SAM effect observed earlier. This does not imply that a school's program participation did not increase its leadership for data-based inquiry. Indeed, earlier survey analyses found that SAM program participation predicted gains on this variable during a two-year period (Talbert et al., 2009). Rather, it means that SAM-cert schools, on average, do not score significantly higher on this measure in 2011 than the non-cert schools. By implication, schools that selected into the SAM program were relatively weak on this leadership approach.

Model II includes the SAM Involvement variable to assess independent effects of LDBI with controls for SAM program participation. Results show significant independent effects for both variables (see third and fourth columns of Table 3). New Visions schools with relatively strong inquiry leadership and/or with teacher teams that participated in SAM beat the odds for getting their students on track to graduate and for college readiness and lessening their likelihood of being seriously off track.

Table 3. Inquiry Practice and Leadership Effects on Student Outcomes

	Model I		Model II (SAM-cert included)	
	% On Track[a]	% Off Track[b]	% On Track[a]	% Off Track[b]
COA	-.02	.13		
LDBI	.30	-.25	.34*	-.26*
% Below Basic Reading	-.05	-.14	-.06	-.08
% Below Basic Math	-.58**	.72***	-.57***	.73***
School Size (n)	-.07	.01	-.31	.19
Proportion Free and Reduced Lunch	-.45*	-.50**	-.31*	.38**
%Asian	-.20	-.10	.09	-.32
%Hispanic	-.17	-.17	-.06	-.24
%Black	.25	-.08	.07	-.34
SAM-cert Involvement			.54**	-.40*
R^2	.64**	.74***	.78***	.80***

Notes: The school N for this analysis is reduced from 45 to 29 since schools with less than a 50 % teacher survey response rate were excluded. School scores are mean teacher scores on each of the two survey scales
* $p < .05$, ** $p < .01$, *** $p < .001$

This finding lends support to SAM's hypothesis that broad leadership of data-based decision making brings more students into the sphere of success – accelerating their progress toward graduation and college readiness. The independent effect of SAM-cert involvement suggests additional benefits of program involvement. According to SAM's theory of change, the benefit is likely to come from the strong role a facilitator/instructor plays in guiding teacher teams' implementation and spread of SAM's inquiry model.

How Inquiry Works and Implementation Challenges

New Visions high schools have implemented collaborative inquiry in varying ways and degrees over the past four years. A small number of schools invested heavily in the SAM model and developed robust inquiry practices and leadership as three or more successive cohorts of teachers participated in the SAM credentialing program (these "mature inquiry" schools were coded "2" on the SAM-cert involvement variable for student outcome analyses). Some schools embraced collaborative inquiry as a vehicle for school improvement and sent a teacher team to the credentialing program during the past 2-3 years (SAM-cert schools coded "1" on the SAM-cert involvement measure). Other schools ranged from nominally satisfying DOE requirements for setting up teacher teams to developing specialized inquiry roles or a school team to embracing inquiry and promoting leadership for data-based improvement.

Longitudinal case studies of four mature-inquiry high schools – two large and two small schools that have had teams in at least three SAM cohorts since 2006-07 – provide evidence of how collaborative inquiry works to bring more students into the school's sphere of success. In each school, teachers changed how they viewed and taught struggling students and the culture shifted toward a norm of teachers working together to diagnose and address needs of all students. These changes came about in schools that, on one hand, had intensive support for developing teachers' inquiry leadership and, on the other hand, present special challenges for changing high school culture (McLaughlin & Talbert, 2007; Siskin, 2011).

Changes in teachers' beliefs and classroom practices

The substantive work of collaborative inquiry and leadership is unique to each school and team. Each identifies particular skill gaps of struggling students and school instructional decision systems that keep them outside the "sphere of success." Yet, collaborative inquiry prompts particular kinds of shifts in teacher perspectives and classroom practices across diverse schools. Our longitudinal research in SAM-cert schools shows 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 classrooms and in observing others, teachers' focus moved from how the curriculum is being taught to what students 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 low-inference transcripts (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 these students' learning needs. Teachers became aware that students had often missed critical segments of content and/or skill instruction provided in earlier grades and that 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 feedback on 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 evidence of particular 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.

Such shifts in teacher perspectives and instructional practices engender and interact with change in team and school culture. For example, individuals' shift in focus from delivering curriculum to diagnosing student learning needs helped teams develop norms of shared accountability for using inquiry to meet the needs of struggling students, and teams' experiences of making a difference for struggling students helped tip the school toward an inquiry culture.

Changes in school culture

Changes in school culture came about as teachers developed collaborative inquiry practices in teams and began to see results. They began to share accountability, hold one another accountable for basing decisions on evidence, and take on leadership roles. These qualitative changes in school culture represent both challenges for, and outcomes of, inquiry-based school reform.

Case studies of schools that invested heavily in this strategy for school improvement suggest that a significant shift in school culture is reached during the third year. This pattern was similar for both large and small high schools in our sample and hints at a developmental arc during which teachers who initially questioned or opposed inquiry as a way to improve instruction for struggling students become convinced that it works.[9] Gradually, teacher leadership supports a school-wide shift toward shared expectations that collaborative inquiry is both a normal part of teachers' work in the school and the most effective way to improve student achievement.

Shift toward shared accountability

As teachers worked in teams to diagnose and respond to the learning needs of struggling students, they began sharing responsibility for the success of all students. They moved from

---

[9] This pattern follows the typical S-curve for innovation adoption in organizations, assuming leadership by 'early adopters'. SAM-cert scaffolds and supports participants' role as early adopters leading school culture change.

thinking about "my" students to "our" students, as well as shifting their attention from successful students to struggling students. This shift pushes against a strong tradition in education to consider teaching the responsibility of individual classroom teachers, as well as against current federal and state system accountability pressures on districts and schools to evaluate and reward individual teachers for their success relative to colleagues.

The shift was supported by SAM-cert because the program holds teams accountable for diagnosing and responding to target student learning needs. Through this practice and sharing a set of students who may or may not be in their classes, teachers developed a sense of mutual accountability that persists. All participants we interviewed explained what a big change it was when they came to feel responsible for the success of all students in the school.

Shift toward norms for evidence-based practice

Teachers developed the habit and norm of using evidence of student performance to evaluate and improve instructional decision-making systems, as well as their classroom practice. Collectively, they moved from: a) relying on individual intuition and past practice to using data to drive their instructional decisions and evaluate student learning; and b) using summative assessments to measure student outcomes to using formative assessments to diagnose student learning needs.

Annual faculty surveys in the four "mature" SAM-cert schools since 2005-06 document these schools' progress toward a culture of evidence-based practice. Figure 3 shows that each school made step increases on our survey scale "Culture of Assessment Use" for the first three years, followed by a plateau or slight decline in recent years (see Appendix C for scale definition). Two are large high schools organized into Small Learning Communities (Schools A and B); and two are small themed schools (Schools C and D).

These schools' inquiry practices today are 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 3). 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 nearly as likely to disagree as to agree with statements that assessments were 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 use of multiple assessments were relatively high.

Through the SAM program, each school made a qualitative shift toward using fine-grained skill assessments to identify and hone in on learning targets for their struggling students, primarily recent immigrant English Learners with weak academic preparation. Significantly, the large schools coupled their inquiry initiative with school redesign into Small Learning Community (SLC). These survey data capture steady and sustained development of inquiry cultures within and across these units in Schools A and B.

Figure 3. Culture of Assessment Use Trends for "Mature SAM" Schools: 2006-2011
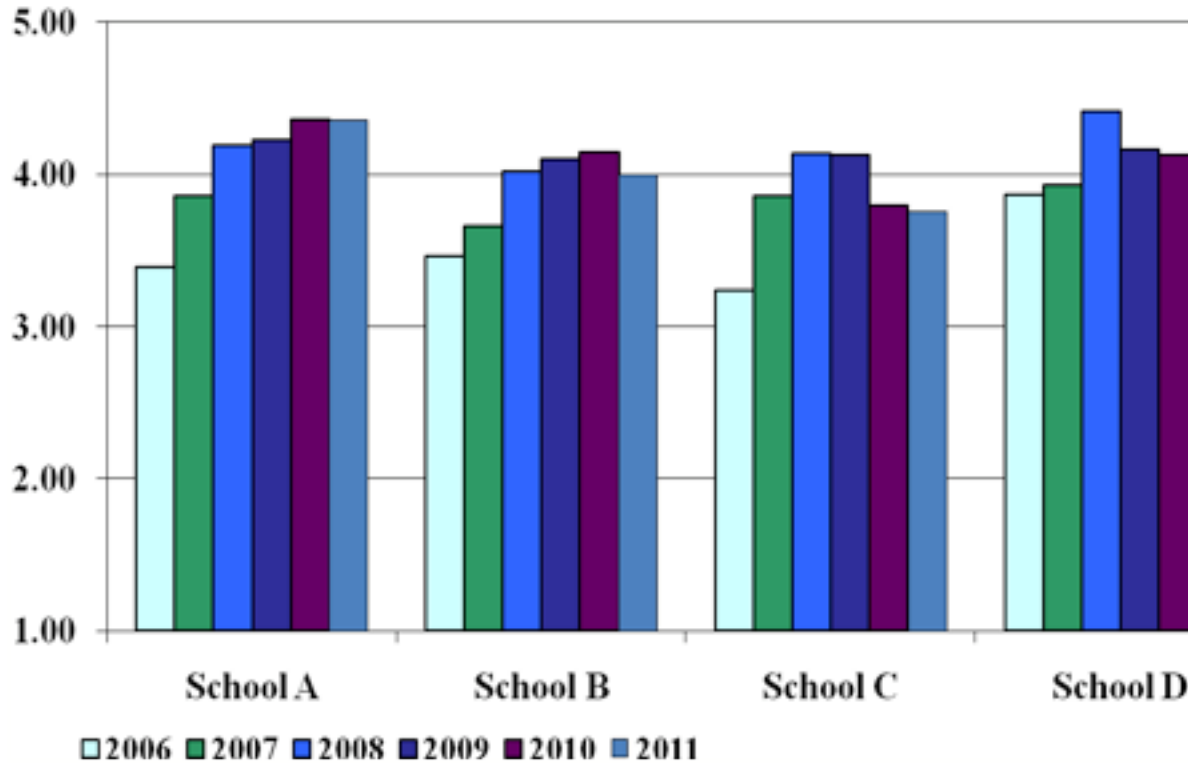
In one of the small high schools (School C), the first SAM-cert team created a new structure for regular school-wide inquiry that yielded a big boost in teachers' use of data on individual students to inform decisions that was sustained for three years, with some decline in the past two years after the teacher leader left to become principal of another school. In the other small school (School D), a school inquiry team including several SAM program graduates took primary responsibility for using data to inform school decisions and to identify and address skill gaps for a target group of students.

Over time in all of these schools, increasing proportions of teachers collaborated on using multiple assessments to focus instruction for struggling students. Much of this work has focused on student skill gaps in literacy, such as identifying the main idea in a text or knowing academic vocabulary. Teacher teams shared instructional responses that have been successful in moving their target students.

Shift toward 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.

In the large restructured schools, successive cohorts of SAM-cert participants took on leadership roles and responsibilities within their SLCs. Inquiry leadership spread and deepened through this process and the active, on-site support of external instructors/facilitators. In one of the schools an Assistant Principal (AP) became a SAM facilitator who teamed up with the external facilitator to lead seminars and support teams' work in the school. Inquiry leadership grew incrementally as teachers were trained to carry out inquiry and lead change in their school and as administrators became invested and skilled in leading inquiry-based reform.

Shift toward on-demand professional development in content instruction

In some schools, teacher teams converged in their efforts to address a specific skill gap prevalent among struggling students, prompting a school-wide instructional response. For example, after three years of SLC-based inquiry work, inquiry leaders in one large high school reached consensus that student writing was a high-leverage skill domain across subject areas and grade levels. Teachers in their SLC teams 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 teams' 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 surfaced 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 classrooms.

In interpreting these results, one issue is whether these schools' success in implementing the SAM model was due to 'selection effects' and 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. So critics may ask whether or not schools that became involved in SAM were already on their way to inquiry-based reform. We cannot rule this out entirely, but we have evidence that these schools had baseline inquiry practices similar to comparison schools. In any case, the mature SAM schools offer evidence that teachers' collaborative inquiry can be a vehicle for changing school culture and bringing more students into a school's sphere of success. The fact that teacher teams in these SAM-cert schools had intensive support from facilitators and that the principal endorsed their work and leadership can be regarded as resources that make a difference for school progress on the initiative, rather than selection bias.

Implementation challenges and resources

All school inquiry teams we studied encountered technical, organizational, and cultural challenges in implementing the model for inquiry-based school reform. Their progress was also not linear. Rather, a team's learning and change was bumpy and cyclical. As teachers moved outside their 'comfort zone' to develop new assessment and instructional practices, they grappled with the tug of old habits and mindsets. Teachers reported moving 'two steps forward and one step back,' needing to 'relearn' a new practice and perspective. They experienced an "aha" only to encounter a new challenge. Some teams got stymied by the road blocks they encountered and

never moved beyond superficial routines of data use; others became highly skilled in using data to continually improve student learning and success.

Evidence of implementation challenges comes from our longitudinal case studies of ten New Visions schools, in which common struggles surfaced in large and small schools, whether or not they had teams participating in the SAM credentialing program. Most fundamental in the developmental trajectory was whether a team and school made 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 shared commitments and a sense of efficacy in diagnosing and addressing student learning needs through instruction and system changes. Survey research in the broader school network provides evidence of resources that matter for progress.

Table 4 summarizes the technical, organizational, and cultural challenges schools faced in implementing the SAM model and the nature of resources that have made a difference for their progress.

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, to use assessment data and student work to identify prevalent skill gaps, and to 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 Data Specialist was a key resource in many schools. Through monthly meetings, Data Specialists learned the ins and outs of the DOE data system and how to analyze periodic assessment data to identify specific skill gaps in student performance. Networking with Data Specialists from other New Visions schools also pointed to effective ways of leading school teams and innovative ways of organizing data.

Nonetheless, 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 seemed counter-intuitive.

In some schools, the SAM or New Visions facilitator 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

Table 4. Implementing SAM: Challenges and Resources.

Challenge for change	Resources for managing the challenge
Technical: Developing inquiry practice	
Using several kinds of data to identify a) skill gaps for under-performing students and b) instructional systems in the school that limit their success	Data-savvy person(s) on team Data system and summaries that include multiple measures and fine-grained data Data Specialist meetings and network SAM model for inquiry and facilitator support in developing the practice
Getting small: honing in on a learning target to focus an instructional response	Skilled facilitator to address resistance and keep focus on 'moving' students on a manageable learning target
Assessing learning outcomes of the instructional response and refining to achieve success with all students	Skilled facilitator to guide design or adoption of pre/post assessment and push the team to persist in getting results
Organizational: Prioritizing inquiry	
Creating and protecting time for collaboration on inquiry	Principal commitment and priority for teachers' collaborative inquiry
Distributing leadership and developing teachers capacity to lead inquiry-based change	Principal delegation of authority to teacher leaders for leading inquiry SAM leadership curriculum and facilitator Coaching
Social-cultural: Shifting norms/ beliefs	
Developing a team and school culture of shared accountability for student success	Administrator focus on results by teacher team; celebration of team success Facilitator support for developing positive team dynamics and leadership skills
Shifting focus from teaching quality to student learning needs and responses	Low inference transcripts (LITs) of classrooms, guided by facilitator Administrator assurance that LITs are not for teacher evaluations
Changing teachers' attribution of student failure and developing sense of efficacy	Evidence of team success in accelerating student achievement and presentations to colleagues

Organizational challenges

An inquiry team needs regular dedicated time for their 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 stalled and teachers saw it as a DOE mandate and took a compliance stance. Principals who made collaborative inquiry a priority in their school understood the model's *principles* and believed it to be an effective vehicle for instructional improvement. In schools where the principal was on board and strategic in involving teacher leaders, teams developed effective inquiry practice over time.

Using two years of data from our annual teacher survey in New Visions network schools, we found a positive statistical effect of Principal Support on change in Team Functioning (see Appendix C for scale definitions).[10] This finding 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 teachers and external and/or internal facilitators. Indeed, broad or distributed leadership was fundamental to inquiry-based reform. Nonetheless, a principal's endorsement of collaborative inquiry as a means to improve student achievement and willingness to share decision authority proved to be essential for progress toward an inquiry school culture.

Our four-year case studies of ten high schools revealed four broad organizational patterns of inquiry implementation that signal differences in the extent of depth and breadth of inquiry development in a school. Two related patterns are characterized by deepening inquiry practice with limited spread. A *specialist* pattern locates responsibility for inquiry with the Data Specialist, who manages student performance data, uses inquiry to identify and diagnose needs of struggling students, and works with students to set goals and determine interventions; teachers are more or less brought into supporting interventions with individual students. A *hub* pattern locates collaborative inquiry primarily in a school team composed of administrator(s) and several teachers from different content areas (the original school Inquiry Team model); other teachers are more or less involved in carrying out interventions with target students. A third, *delegation* pattern is characterized by breadth with limited depth of inquiry, in which all teachers meet regularly in grade-level teams and examine student performance data to identify gaps and decide on interventions for struggling students but lack leadership to develop rigor and mutual accountability for results. The fourth, *leadership development* pattern is characterized by broad and deepening inquiry across the school, reflecting successive training of teachers in how to conduct and lead inquiry with colleagues to continuously improve student success. As noted, the SAM-cert program has been pivotal in the large restructured high schools we studied.

These patterns express school administrators' differing priorities and approaches to implementing collaborative inquiry in their school. They also reflect and influence teachers'

---

[10] School scores on the Principal Support scale predicted their scores on the Inquiry Team Functioning scale after the prior year's Team Functioning score was controlled. Coefficients were .78 (p=.01) and .35 (p=.05), respectively.

understanding of, and capacity to use, the SAM inquiry model to move students on a high-leverage learning target. Pivotal to the fourth pattern of developing inquiry leadership was a school's partnering with a skilled inquiry facilitator who worked with a team of teacher leaders to develop their inquiry skills and habits of mind and readiness to lead inquiry across the school. The facilitator's role was particularly crucial in nurturing shifts in teachers' beliefs and culture essential to effective inquiry practice and to success in leading school change.

Professional culture challenges

Leaders of inquiry-based school reform face opposition and obstacles that stem from long-standing norms in teaching. Although some schools have reform histories and founding cultures that depart from traditional norms and conventions, all are challenged to address professional beliefs and habits that constrain their progress on inquiry to achieve success with struggling students in their school.

The 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 in the schools we studied held this focus when teachers resisted opening their classroom and their students' work to colleagues. Administrators pushed for teachers' shared accountability by focusing on grade-level, department, and SLC performance in reviewing school progress – promoting a view of teaching quality as a collective, rather than individual, matter.

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 pointed to the practice of low-inference transcripts (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," "interpret" – realizing that the students could not comprehend such directions or access instructional content.

School administrators and team facilitators play important roles in implementing LITs, since teachers sitting in on their peers' classes and documenting dialogue is counter-normative in teaching and threatening to many teachers. We documented considerable resistance and struggle with this practice in the early years of SAM, and program refinements provided more support for teachers' scripting and fewer observation demands. Overcoming peer resistance depended on administrators making clear that the classroom observations were not being used for purposes of teacher evaluations. Facilitators supported teachers in developing the skill of scripting classrooms verbatim so that student experiences became salient and accessible for developing instructional responses. Teachers found protocols for analyzing the LIT helpful for tracking the progress of their instructional responses. For example, a team that focused its intervention on developing students' academic language used LITs to assess target students' use of the terms in their classes.

Developing teachers' sense of efficacy or confidence that they can meet student learning needs presents a significant challenge for culture change. Convention has it that students'

academic failure is rooted in difficult family conditions or traits like 'laziness' or personal troubles that get them off track for success. Such accounts of poor student outcomes were common among teachers in teams that lacked a solid design and support for inquiry. Related were assumptions that learning in content areas 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 succeeded in improving target students' performance on a specific skill or learning target, 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.

In order to achieve this success with students, however, a school team needed support for tackling the technical, organizational, and cultural challenges for change. As noted in Table 5, a primary resource was a facilitator who deeply understands principles underlying the inquiry model. The facilitator guided the team to develop inquiry practice and to address challenges of shifting their beliefs and leading culture change in the school.

Further investigation of the facilitator role in school progress on inquiry-based reform helps to interpret SAM-cert schools' relatively strong performance on student outcomes. A key difference in resources available to SAM-cert and non-cert schools was the intensity of facilitator support the teams got. As described earlier, SAM instructors used a tested and refined curriculum and protocols to scaffold their team's work on inquiry and leadership, met with teams in weekly seminars to guide and give feedback on their inquiry work, and coached individuals in using inquiry in their classroom and in leading change with colleagues. Further, they themselves formed an inquiry team of facilitators to learn to improve their success with school teams. Insight into the challenges facilitators face in effectively guiding inquiry practice and school culture change helps explain why investing in their preparation and ongoing learning is important to schools' success.

Challenges and Resources for Facilitating Implementation

Research documents the importance of a facilitator in nurturing the development of inquiry practices in teacher teams. Yet little attention has been paid to what it takes for a facilitator to be effective. What challenges does a facilitator face in working with a school and what resources make a difference for their success in facilitating school progress on inquiry-based reform?

Although the LDFs and SAM facilitators shared the goal of developing team inquiry practices and spreading them across a school, they had different kinds of resources to draw on for achieving this objective. These differences surfaced fundamental challenges of the role and what matters. For example, LDFs and SAM instructors both grappled with adapting their work to teams and schools that differed in inquiry readiness. Yet they had different resources to draw upon, which may or may not matter for any given challenge. This analysis helps to pinpoint resources for developing high-quality facilitation for inquiry-based school reform.

Fundamentally, facilitators helped teams grapple with the challenges they faced in their efforts to implement the inquiry model and develop a school culture of shared accountability, evidence use, and distributed leadership. Table 5 highlights the main challenges facilitators faced in doing so and what they identified as supporting their success in meeting each challenge.

Technical challenges and resources

Inquiry facilitators needed to *establish a clear focus and priorities for their work* with school inquiry teams. Although the DOE and New Visions network adopted the SAM model and developed data systems and outputs to support school-based inquiry, facilitators struggled to make the vision real for school teams. Many lacked training to facilitate deep inquiry and drew upon whatever supports they had to just manage interactions with the school. In the words of one LDF: "I took advantage of the coaches, the ex-principals, former principal coach…you just brought in whatever you could. Because…you don't have that expertise." Although they had access to a variety of inquiry tools, they often lacked an understanding of what inquiry tools can do and how to use them in their work with school inquiry teams.

In turn, facilitators grappled with how to *support teams' use of tools for inquiry and assessments of student learning*. Many found available inquiry tools to be lacking and/or themselves to be limited in necessary skills to support teams' use of them, particularly the use of formative assessments to measure student learning. As one facilitator put it: "So teachers did definitely select a strategy, they definitely implemented it in the classroom, they had a goal in terms of an academic outcome. But the interim assessment piece and the ability to measure real changes along the way were not there…because they don't always have the tools." In order to move the inquiry work forward, LDFs supported teachers in designing their own assessment tools, "but the inquiry process itself is hampered to some degree by not really having a great way to measure progress."

Table 5. Facilitating SAM Implementation: Challenges and Resources

Challenge for facilitation	Resources for managing the challenge
Technical: Developing team inquiry skills	
Learning core principles of the SAM inquiry model; establishing focus and priorities for a particular team	Guidelines and pillars for team inquiry Facilitator PLC: debriefing practice and refining strategies
Introducing and scaffolding team inquiry tasks and use of protocol/tools	Training in use of data and inquiry tools (ARIS, Data Tracker, LITs) SAM curriculum Learning from own and colleagues' experience
Developing a team's skills in designing and using assessments	Assessment design training and access to formative assessments for learning targets
Organizational: Navigating school context	
Negotiating access and role as facilitator of inquiry team practice	Principal priority for team inquiry and authorization of facilitator role Common planning time for inquiry teams
Customizing work with diverse schools (e.g., grade level, size, student population, principal support level)	Facilitator PLC: problem solving around priorities and practices effective with particular school types
Gearing facilitation to inquiry team readiness/ developmental stage	Developmental view of inquiry team practice Facilitator PLC: dialogue around practices suitable at team readiness levels
Cultural: Addressing norms and beliefs that inhibit inquiry progress	
Nurturing trust, collaboration, and shared responsibility for student success in a team	Policy/ communication focused on student learning (vs. teacher quality) Specialized roles for team members to adopt SAM team-building protocol
Challenging team members' attribution of student failure to external factors; shifting focus from teaching as 'delivering a curriculum' to 'ensuring student learning'	Inquiry guideline to "stay small" so that a team can experience success in moving students Facilitator PLC: developing ways of pushing new ways of thinking about instruction
Helping teacher leaders develop skills in spreading inquiry in their school and addressing school-specific challenges	SAM readings on organization culture change, protocol for analyzing colleague resistance, guidelines for inquiry practice, facilitator protocol for modeling, fishbowls, role play

SAM facilitators addressed the same set of problems in moving school teams on inquiry. However, their work was scaffolded by the SAM curriculum that set stages and steps for the team to move forward. Their challenges centered on using the curriculum to focus teacher teams' inquiry with their own students in their schools. Many SAM-cert teams struggled with creating measurable and granular learning targets, according to the facilitators and teams themselves. When confronted with issues like the fact that the team was "flailing with their larger purpose/goal" in designing specific target skills, facilitators turned to each other and their trainers for support. For example, in one instance, when "feeling a bit undone before the session – didn't feel as though I really knew my objectives," one facilitator reported that observing her trainer's coaching with a fellow facilitator around the same issue helped deepen her understanding of the coaching task.

LDFs and SAM facilitators alike reported that sharing, learning from, and debriefing with other facilitators helped them interpret their responsibilities and support their work with the schools. Opportunities to work together in small groups helped them learn and understand their roles and how to make use of available tools. In the case of SAM facilitators, weekly seminars supported them in "norming" their own practice and team leadership while developing individual styles.

Organizational challenges and supports

Facilitators needed to *negotiate access and their role as inquiry facilitator* in a school. This challenge was greatest for LDFs whose work with inquiry teams was just one facet of their work. They faced challenges of getting school leaders' buy-in to inquiry, ensuring time for collaborative inquiry, and negotiating their role in facilitating inquiry team work. Some had difficulty getting access to a school at all, let alone to support the inquiry team work. As one explained: "I was never really able to make a good connection with the school leader…I didn't really have adequate access to the school. So *all* the work I've done has really been *me* pulling information and doing my updates on how they're doing." Another expressed frustration in trying to get traction in a school: "I thought I was really beginning to move the place. But I feel that I'm ending up in the same place as my predecessors [who couldn't get access the previous year]." Turnover in LDFs working with a school was one source of the problem of negotiating relationships with schools, since building trust during the first year of working with a school was key to moving forward in a role as change agent.[11]

The challenges facilitators faced in *customizing work with diverse schools* includes accommodating different principal stances on inquiry and being strategic about leveraging change. Negotiating support for an inquiry agenda was vital. One LDF described efforts to work with a principal and faculty resistant to implementing inquiry:

> The principal told me, "My teachers don't want to hear anything more about data." And it took me from September until two weeks ago [in April] to actually get him to touch a real child's permanent record…I felt very much that they were totally unwilling to engage.

---

[11] Our 2010 report provides an analysis of LDF turnover and its effect on working relationships with a school For example, 75% of LDFs who worked with schools for two years report playing an active role in supporting their inquiry work, compared to 44% of those who moved from another school and 50% who were new to New Visions.

Another facilitator encountered difficulties in working with assistant principals: "I would have liked to have been part of the talk about planning for inquiry. But there's a resistance, more by [the] assistant principals." Another said:

> I'll just say what my one big learning is: if the principal does not participate—it does not happen…or it may happen, but it won't have impact on the school…So, to me, it's about saying…'Everybody is participating in this. This is the way we work. It's not about a little project, a little special thing aside. It's really the core of the way we work.'

SAM-cert schools met at least a threshold of principal support for a teacher team's participation in the credentialing program. Yet, SAM facilitators worked with schools that varied widely in the extent to which the principal was knowledgeable and actively supportive of inquiry as a vehicle for change. In instances where the administration was initially neutral or skeptical regarding inquiry as an approach to improve student learning, facilitators negotiated entrée and their role in supporting inquiry, including boundaries for their respective leadership roles in the seminars.

Facilitators constantly faced the *challenge of adapting their supports to a team's readiness*. LDFs made judgment calls about a school's readiness to take on inquiry in a serious way and in some cases backed off of the work if other school priorities seemed more pressing. For example, Transfer Schools were generally given different priorities for change. A notable exception was a transfer school in our case study sample that made great strides on inquiry. The LDF worked with the school inquiry team and encouraged two teachers to join the SAM program to advance their inquiry practice and leadership development in the school. This two-person team stepped up to lead inquiry in the school and made good progress within the first year and a half.

SAM facilitators worked in their weekly meetings on challenges of adapting their practice to the developmental needs of particular teams in their cohort. One strategy they developed was structuring "group time" at the beginning of each seminar in order to check in with each team to provide differentiated instruction and support. In the words of a SAM facilitator: "group time…gave me the opportunity to check in with the teams – one was on track…[and] the other needed more coaching to get to understand what I was asking of them." Facilitators' weekly meetings regularly took up issues of how to address places where particular teacher teams were struggling and in the next session checked in on their progress and next steps.

Cultural challenges and supports

Facilitators aimed to *nurture team trust and collaboration*, often among teachers who had never worked with one another or who had a history of contentious working relationships. Because inquiry teams needed to open up their practice with one another, a condition of mutual respect and trust was essential; for many teachers, sharing their own work and struggles with school colleagues proved to be threatening.

As schools required multiple teacher inquiry teams across grade levels and content areas, facilitators faced the challenge of developing teachers' facilitation skills in order to create conditions of trust in their collegial teams. Facilitators also played a role in establishing school guidelines for team roles to help build ownership of the work. As one LDF explained:

> One of my lessons learned this year is that it makes a big difference how you launch those teams. Who are they accountable to? Who on the team is accountable? What kind of preparation do they have? I feel like we did a lot of things really well. For example, in the beginning of the year we did develop very tight protocols and agendas and had everybody doing the same thing. And we had administration representation on the teams. Not necessarily to be the leader; but to be the link, the liaison, and to hold the team accountable, and to bring back the learning. Right? As we progressed through the year the teams became more independent. In some cases we were able to not even have a consistent administrative presence on the teams. They began to take ownership. And they also began to differentiate according to the challenges that they had. So we didn't have to script agendas for them anymore. Some of them had taken off and done *wonderful* work. And they really have a cohesive sort of team spirit. They create their own agendas. They do their own data analysis.

Because SAM-cert explicitly aimed to develop participating teams' skills in leading colleagues in inquiry-based reform, the program designed various tools and strategies to develop positive team dynamics. For example, exercises and readings in the first curriculum module addressed issues of negative dynamics with colleagues and ways of developing a complementary division of labor for team work. Also, the facilitators regularly interacted with teams about their internal dynamics. During "team work" time in seminars, facilitators prompted participants to reflect on individuals' roles and responsibilities on their teams, how the dynamic evolved over time, and ways in which their collaboration could be strengthened. In addressing teams' different needs and developmental stages, one facilitator divided participant teams/schools into three levels: high touch ones that needed a great deal of support, mid touch ones that needed some support, and low touch ones that functioned independently and required only periodic check-ins based on need. Further, SAM facilitators led sessions with multiple teams that addressed distinct team dynamics and leadership roles. They found that "critical friends groups" – which they role played during their weekly facilitator-only meetings, rotating roles of presenter and facilitator – were a valuable resource for addressing team dynamics and developing interventions.

Apart from developing positive team dynamics, facilitators aimed to *disrupt teacher beliefs that inhibit student success*. They played a crucial role in moving individuals and teams beyond their comfort zone – challenging beliefs and habits of mind that typically "excuse" student failure. As described earlier, external attributions of failure and a view of teaching as curriculum delivery, were chief among them.

Facilitators prompted team members to break their habits of mind. In our annual surveys, the majority of teachers agreed that a facilitator "pushes us to think in new ways" and "raises good questions that move forward our thinking about the work." When present during inquiry team meetings, a facilitator pushed conversations away from a student's family circumstances and toward his/her academic skills and gaps, for example, and raised questions

about where in the curriculum a struggling student can address a particular skill gap revealed by assessments.

SAM included three "anchors" of facilitation to shift a team's habits of mind and routines that inhibited their success with struggling students. The guidelines helped facilitators meet the challenge of re-culturing a team. First: "provoke and support learning" by both creating enough tension to prompt team members' learning and by not acquiescing to their desire to remain in or return to their "comfort zone." Second: "keep the focus on results" by making the team's targets clear and public, making sure that the targets are owned by the team, and helping them manage distractions. Third: "ensure timely, honest, and actionable feedback so that teams can learn through their inquiry work and act on evidence showing areas for improvement. SAM facilitators held themselves accountable for following these core principles for practice, and they guided participants to use them in leading inquiry teams in their schools.

Finally, in supporting transition to the final phase of SAM's model for inquiry-based school reform, facilitators were challenged to *help inquiry team leaders develop an inquiry culture in the school*. These challenges centered on problems of developing change leadership, a problem space that is the focus of roughly a third of the SAM curriculum and assignments. They also centered on avoiding a compliance stance on inquiry, given the DOE mandate for schools to develop collaborative inquiry teams (see Talbert, 2011).

LDFs faced the challenge of getting schools past the frame of the DOE's mandate for collaborative inquiry. As one described: "I think all of that also sits in [the] context of a mandate, of a thing that's now 'You have to do this.' And 'inquiry' has become not the best word in some schools."

Still, an inquiry stance grew in schools where the teacher teams shared evidence of improved student outcomes with their colleagues. A clear focus on student learning and success with struggling students was compelling to most teachers. One facilitator commented that teachers have gradually shifted in their stance on inquiry:

> There are folks that absolutely think that 'I'm the teacher here, and I've been teaching, and you can't really tell me—because I'm a perfectly good teacher.' So there's been a little bit of a culture shift in terms of some of the [that] grade teachers. But I think they've come around. I mean they may not totally be on the same page, I would say, but they really do care very much about the kids and they understand that this method—the inquiry process—has really moved the kids. I mean that's just undeniable.

Perhaps the greatest challenge SAM facilitators faced was developing team members' skills in facilitating inquiry with their colleagues –prompting other teachers to rethink their assumptions about student learning and instruction. The goal was that SAM graduates would be able not only to manage resistance and motivate colleagues toward inquiry, but also lead individual teachers and teams to investigate their practice in terms of student learning. This presented a big charge for individuals who participate in their school's teaching culture.

SAM's coaching protocol, developed and refined by the facilitators, was a response to this challenge. The protocol prompted SAM participants to work with a teacher colleague to brainstorm elements of a lesson and learning objectives, then to script the lesson as evidence of what happened, and finally to debrief with the teacher in ways that would stimulate new thinking and plans. In order to assess participants' coaching practice, SAM facilitators scripted and analyzed the team members' coaching sessions. It became apparent that participants had great difficulty adopting a coaching stance towards their colleagues. Facilitators' analysis of the relationship between a teacher's lesson objectives and classroom scripts identified inconsistencies. Yet, the coaching scripts revealed that SAM participants typically did not address the inconsistencies. Either they did not detect them or they preferred to maintain a congenial relationship with their colleague rather than create disequilibrium. Based on this information and plausible explanations, the facilitators both refined the protocols and their coaching of SAM teams on leading change in their schools.

This example of problem solving around facilitator practice illustrates how SAM facilitators' weekly sessions featured inquiry to continuously improve their practice in working with school teams. LDFs and SAM facilitators alike pointed to dialogue and collaboration with colleagues – an inquiry team or professional learning community (PLC) of their own – as the greatest resource for their learning and success in meeting the challenges of their work.

* * * * *

Our research on schools implementing the SAM inquiry model, as well as broader literature on professional learning communities, suggest that external coaches or facilitators are key agents in leveraging and supporting change in a school's instructional culture (Gallimore, Emerling, Saunders & Goldenberg, 2009; McLaughlin & Talbert, 2006, 2007). This is because routines and habits of mind in teaching, as well as instructional policy, tend to inhibit change – especially toward collaboration in using evidence to improve instruction (Coburn & Talbert, 2006; Ingram, Louis, & Schroeder, 2004). The goal of accelerating learning for students who are likely to fail goes against the position that such students are victims of poverty and dysfunctional families and schools cannot make much difference. A skilled and trusted facilitator can jar educators' thinking in ways that enable them to move past prior assumptions and tackle the challenge of ensuring all students' success. Yet little research informs our understanding of what and how facilitators learn to do this work effectively. This analysis is a step in that direction.

Promise and Prospects of SAM as Vehicle for Inquiry-based School Reform

Lessons from schools' implementation of SAM over the past four years contribute to a growing literature on inquiry-based school improvement. SAM's highly-specified model for collaborative inquiry and teacher leadership of school change set it apart from typical approaches to inquiry-based reform. Schools that invested heavily in implementing SAM developed an inquiry culture and expanded the sphere of student success. These schools involved successive teacher teams in the certification program and benefited from its implementation supports – curriculum and standards, intensive facilitation, and network of participating school teams.[12]

Evidence that SAM's theory of school change works to develop a collaborative inquiry culture and improve student success argues for sustaining efforts to spread and deepen the model in NYC, and beyond. Yet, challenges of implementing SAM were considerable, and collaborative inquiry failed to get traction in schools that gave it low priority and/or lacked strong team facilitation.

Here we take stock of our evaluation findings, considering SAM's promise for continuously improving student success and prospects for spreading and deepening its use in NYC schools and beyond.

Promise: Outcomes and capacities developed through SAM

Schools with sustained teacher involvement in the SAM credentialing program had relatively strong student outcomes. On average, these New Visions high schools outperformed comparison schools with comparable student populations. They had significantly larger proportions of students on track for graduation and college readiness and smaller proportions significantly off track for graduation by their junior year.

This does not imply that the comparison schools did not implement SAM (though they varied widely) or improve their students' outcomes.[13] Rather, it means that schools that invested most heavily in developing teachers' capacity to implement and lead SAM did better in getting students on track for graduation and college readiness. Investing in collaborative inquiry proved to be a more robust and cost-effective strategy than common alternatives.

SAM's theory of change posits that inquiry leadership and culture shifts are the source of improved student outcomes and that collaborative inquiry is the engine of change. Evidence from the mature SAM schools we followed for four years supports the argument. Consistent across

---

[12] As noted, the certification facet of the program appears to be less important for implementation than these resources. Indeed, one of the large high schools most successful in implementing SAM did not involve teachers in the credentialing part of the program, yet teacher teams participated in all other facets of the program and became instructors/ leaders with successive cohorts of their colleagues.

[13] Indeed, MDRC's analysis of student outcomes in New York City documents a general trend of improved student outcomes across NYC schools. The study estimated effects of the Children First initiative, including its Collaborative Inquiry initiative modeled after SAM, and found improved student test performance and graduation rates (Kemple, 2011).

these schools were shifts in teachers' assumptions about struggling students' abilities to succeed and their ability to make a difference shifted through inquiry. The teams developed shared accountability and a sense of collective efficacy, or the belief that their work together will improve the achievement of their schools' struggling students. These findings dovetail with prior research on outcomes of collaborative inquiry in teacher teams (cf. Gallimore et al., 2009).

SAM's promise goes beyond the outcomes observed to date. Schools with a collaborative culture have the capacity to respond to new national standards for student learning. Common Core Standards (CCS), embraced by New York and most other states, call upon educators at all levels of the system to prepare students for widely-endorsed 21st century standards for college and career readiness. Schools that have a collaborative inquiry culture are better prepared to respond. Teacher teams have learned to use formative assessment data to diagnose and address student learning needs. Absent teachers' ability to address their students' skill gaps, new standards for higher-order learning are bound to leave students farther behind.

Further, SAM created a pipeline of certified administrators and program graduates who are prepared to become APs and principals with deep knowledge of what it takes to use inquiry to improve student achievement. They are ready to step into administrator positions and are prepared to lead inquiry-based school reform.

A broader frame on the promises of SAM takes into account alternatives for improving students' success. Typically, teacher professional development in content instruction is seen as the route to improved teaching and learning. Yet, research shows that teacher learning and change in instructional practice is a slow process. Teachers' collaborative inquiry is a more robust and cost-effective strategy for addressing the needs of struggling students.

Prospects: Implementation resources and sustainability

Prospects for sustaining and spreading SAM hinge on the appetite for this approach to school improvement and adequacy of implementation resources. Most fundamentally, implementation depends upon school administrators' commitment to inquiry-based school reform and facilitators deeply grounded in SAM principles. What is the likelihood that increasing proportions of school principals will come to see collaborative inquiry as their primary improvement strategy? And, what are the prospects for growing the pool of skilled SAM facilitators?

The SAM credentialing program has contributed a pool of prospective school administrators skilled in and committed to collaborative inquiry. Its graduates implemented the model and understand how and why it works to improve student success. Numbers currently in administrative positions are not large, since many have remained in teaching and school leadership. However, they do constitute a pipeline for Assistant Principals and Principals prepared to invest heavily in implementing SAM.

The program also developed a pool of individuals skilled in facilitating team inquiry. They include SAM instructors who worked with each of the five cohorts of school teams. They also include school administrators who partnered with the external facilitators in leading teams'

collaborative inquiry. Along with SAM graduates, these experienced facilitators are prepared to work effectively with schools in implementing inquiry-based reform.

Broader principal commitment and investment in facilitator development will depend upon system and school leaders' access to evidence that SAM is a cost-effective approach to improving student achievement. NYC's school networks are the primary context in which principals might learn about and become committed to SAM. Thus, network leaders are pivotal to its future. The potential for SAM to become a more powerful reform vehicle depends upon whether and how networks privilege the reform strategy. Do they share evidence of its effectiveness and model evidence-based decision making? Or, conversely, do they promote alternative improvement strategies, such as teacher professional development in content instruction?

The larger system context pulls in both directions. On one hand, NYC's Department of Education continues to promote collaborative inquiry as a strategy for developing school capacity to improve student achievement. Schools are held accountable for establishing inquiry teams, and their Progress Report rating is influenced by evidence of teachers' collaborative inquiry. On the other hand, the mayor and state are promoting an approach to teacher evaluation that could well undermine collaborative inquiry. A system that rates and rewards individual teachers on the bases of their own students' test schools could pit teachers against one another or at least inhibit their collaboration to improve all students' achievement. Despite this policy incoherence and issues over the reliability of value-added assessments, it seems likely that NYC will move forward on this model. Federal Race to the Top regulations make funding contingent on a state's use of value-added teacher evaluation and thus implicitly pushes against teacher collaboration. The national move in this direction is likely to inhibit educators' appetite and capacity for collaborative inquiry and shared accountability.

References

Coburn, C.E., & Talbert, J.E. (2006). Conceptions of evidence use in school districts: Mapping the terrain. *American Journal of Education, 112 (4),* 469-496.

DuFour, R., Eaker, R., & DuFour, R. (2005). Closing the knowing-doing gap. In R. DuFour, R. Eaker & R. DuFour (Eds.), *On common ground: The power of professional learning communities* (pp. 225-254). Bloomington, IN: National Education Service.

Fairchild, S., Gunton, B., Donohue, B., Berry, C., Genn, R. & Knevals, J (2011). *Student progress to graduation in New York City high schools: Part I: Core components.* New York, NY: New Visions for Public Schools.

Fullan, M. (2005). *Leadership and sustainability: System thinkers in action.* Thousand Oaks: Corwin Press.

Gallimore, R., B. A. Ermeling, W. M. Saunders, C. Goldenberg (2009). Moving the learning of teaching closer to practice: Teacher education implications of school-based inquiry teams. *Elementary School Journal,*109 (5): 537-553.

Heifetz, R.A., & Linsky, M. (2002). *Leadership on the line: Staying alive through the dangers of leading.* Cambridge, MA: Harvard Business School Press.

Ingram, D., Louis, K.S., & Schroeder, R.G. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *Teachers College Record, 106(6),* 1258-1287.

Kemple, J. J. (2011). Children First and Student Outcomes: 2003-2010. In J. A. O'Day, C. S. Bitter & L. M. Gomez (Eds), *Education reform in New York City: Ambitious change in the nation's most complex school system* (pp.255-292).Cambridge, MA: Harvard Education Press.

McLaughlin, M., & Talbert, J.E. (2006). *Building school-based teacher learning communities.* New York: Teachers College Press.

McLaughlin, M., & Talbert, J.E. (2007). Building professional learning communities in high schools: Challenges and promising practices. In L. Stoll & K.S. Louis (Eds.), *Professional learning communities: Divergence, depth, and dilemmas* (151-165). New York: Open University Press.

Popham, W. J. (2001). The Truth About Testing: An Educator's Call to Action. American Society of Curriculum and Development.

Scharff, N., DeAngelis, D., & Talbert, J. (2010). Starting Small for Big School Improvement. *Principal Leadership,* April.

Siskin, L. S. (2011). Changing contexts and the challenges of high school reform in New York City. In J. A. O'Day, C. S. Bitter & L. M. Gomez (Eds), *Education reform in New York City: Ambitious change in the nation's most complex school system* (Pp 181-198). Cambridge, MA: Harvard Education Press.

Talbert, J E. (2011). Collaborative inquiry to expand student achievement. In J. A. O'Day, C. S. Bitter & L. M. Gomez (Eds), *Education reform in New York City: Ambitious change in the nation's most complex school system* (Pp 131-156). Cambridge, MA: Harvard Education Press.

Talbert, J.E., Mileva, L., McLaughlin, M., & Cor, M.K. (2010). *Leadership Development and School Reform through the Scaffolded Apprenticeship Model (SAM). Case Study: Hillcrest High School*. Stanford University: Center for Research on the Context of Teaching.

Talbert, J.E. with Mileva, L., McLaughlin, M., Schoener, J., Cor, M.K., Chen, P., & Lin, W. (2009). *Leadership Development and School Reform through the Scaffolded Apprenticeship Model (SAM)*. Stanford University: Center for Research on the Context of Teaching.

Appendix A. *Correlations among School Variables for Student Outcome Analysis*

	% Below Basic Reading	% Below Basic Math	School Size	% on Free or Reduced Lunch	% Asian	% Hispanic	% Black
SAM-cert Involvement	.076	-.073	.538[**]	-.275	.007	.101	-.293
% Below Basic Reading		.362[*]	-.093	.198	-.189	.556[**]	-.317[*]
% Below Basic Math			-.002	.141	-.061	.072	.038
School Size				-.168	.131	-.193	-.129
% Free or Reduced Lunch					-.192	.344[*]	-.055
% Asian						-.319[*]	-.450[**]
% Hispanic							-.635[**]

Appendix B. *SAM-cert and Non-cert School Means on Matching Variables*

	SCI	% Below Basic Reading	% Below Basic Math	% FRL	School Size (N)	% Asian	% Hisp	% Black	% White
Overall (N = 26)									
Mean	.52	9	13	65	330	14	49	29	7
SD	.16	11	8	22	349	16	20	21	13
SAM-cert Schools (N = 13)									
Mean	.53	11	13	62	461	16	52	20	12
SD	.17	14	8	25	448	15	21	16	17
Non-cert Schools (N = 13)									
Mean	.51	7	12	68	199	12	47	39	2
SD	.15	5	9	19	124	18	18	21	3

Appendix C. *Survey Scale Definitions*

       Survey scales were developed to measure school culture outcomes and implementation conditions in the SAM theory of change. Survey items that make up each scale are reported below. Teachers surveyed in 2011 included members of their school's Inquiry Team in 2008 and/or 2009. Over the four-year evaluation, survey respondents included 187 teachers from 57 schools in 2008; 291 teachers from 71 schools in 2009; 164 teachers from 58 schools in 2010; and 264 teachers from 62 schools in 2011.Response rates ranged from 45% to 64% across all New Visions schools, with an average of over 80% each year in case study schools.

       This report includes results of analyses for selected survey scales. Survey items that make up each scale used 5-pt Likert scales from 1 ("Strongly Disagree") to 5 ("Strongly Agree"). Principal components analysis techniques were used to identify items that load on a common factor. Alpha coefficients indicate the internal consistency of a scale and are reported here for 2009 data. Scale scores equally weigh component items and are reported as means (1-5).

INQUIRY CULTURE

➢ *Culture of Assessment Use (2 items. Alpha = .82)*

    *How well does each of these statements describe **how teachers work together in your school or SLC** (if you teach in a large high school divided into SLCs)?*

* We use a variety of assessment strategies to measure student progress
* We use assessment data to evaluate our curriculum and instructional practices

➢ *Leadership for Data-based Improvement (4 items. Alpha = .91)*

* Aligns every action with improvement of student outcomes
* Uses data to identify patterns to inform decision making
* Uses objective evidence to identify, frame and solve problems
* Uses data to evaluate the effectiveness of decisions

INQUIRY TEAM FUNCTIONING

➢ *Team Functioning (4 it*ems, Alpha = .94)

* End team meetings with clear and specific understandings of actions to be taken, timelines, and distribution of responsibility
* Work as a group to equitably distribute the workload
* Know what each of us is working on and how this contributes to the group
* Leave meetings confident that we all are committed to the decisions agreed upon, even if there was initial disagreement

INQUIRY TEAM SUPPORT

➢ *Principal Support of Inquiry Team* **(3 items. Alpha = .88)**

To what extent does each of the following statements capture your **Inquiry Team's** *experience with principal support over the past year?*

* Principal establishes conditions for trust and open communication
* Principal actively supports our risk-taking
* Principal uses authority to push our learning in the service of target students and targeted learning goals

➢ *Facilitator Support of Inquiry Team* **(10 items. Alpha = .96)**

To what extent does each of the following statements capture your **Inquiry Team's** *experience with your facilitator over the past year?* [Note: "facilitator" refers to the New Visions LDF working with your team or, if you participate in the SAM program, to your instructor.]

* Facilitator conveys clear objectives and expectations for our work
* Facilitator creates structures for feedback and self-assessment on our behavior
* Facilitator elicits, respects, and incorporates multiple voices and perspectives
* Facilitator pushes us to think in new ways
* Facilitator helps us shift direction and make corrections when we reach an impasse
* Facilitator raises good questions that move forward our thinking about the work
* Facilitator pushes our learning even when it causes discomfort or anxiety
* Facilitator knows when not to push and how to contain anxiety
* Facilitator continually reinforces the core ideas of inquiry to widen the sphere of student success in the school key concepts in inquiry
* Facilitator holds us to the performance standards for inquiry teams, specifically, moving the students