

**Center for Research on the Context of Teaching
Stanford University**



**Leadership Development and School Reform through the
Scaffolded Apprenticeship Model (SAM)**

Joan E. Talbert
with
Lambrina Mileva
Milbrey McLaughlin
John Schoener
M. Ken Cor
Pai-rou Chen
Wendy Lin

October 2009

The SAM evaluation is being conducted under the auspices of New Visions for Public Schools. Primary funding for the SAM II evaluation came from the U.S. Department of Education's School Leadership program through the New York City Leadership Project. The SAM III-IV evaluation is funded by the Carnegie Corporation. Opinions expressed here do not necessarily represent those of New Visions or of its funders. We thank Ronald Chalusian and Beverly Donohue of New Visions and Nell Scharff and Liz Gewirtzman of Baruch College for their ongoing support of the evaluation and for comments on an earlier draft of this report. We are grateful to the principals and inquiry team members of all SAM II and New Visions' PSO schools for the time and thought they have contributed to our annual surveys, interviews, and focus groups. Special thanks to the Marble Hill SAM teams whose work over four years is featured in this report.

Table of Contents

Introduction	2
SAM’s Rationale and Research Grounding	3
Limitations of Traditional Credential Programs	3
Challenges of Inquiry-based School Reform	4
SAM’s Theory of School Change and Design	5
The Inquiry Team	5
Tasks to Guide Inquiry Practice	6
Tools to Frame and Support SAM Teams’ Work	6
Facilitators	7
Research Methods, Data, and Analyses	10
Examining Evidence of SAM’s Effectiveness	11
Assessing School Culture and Leadership Outcomes	12
Assessing Principal and Facilitator Effects on Inquiry Progress	12
Case Illustration of How SAM Works Up Close	12
Evidence of SAM’s Effectiveness	12
SAM as an Administrator Credentialing Program: Pipeline Outcomes	13
SAM as a School Reform Model: Improved Student Achievement	15
School Culture and Leadership Outcomes	17
SAM Schools Moved Toward a Culture of Inquiry	17
SAM Teams Led Inquiry-based Reform in the School	18
Principals Make a Difference for Team Progress	21
Facilitators Make a Difference for Team Progress	22
How SAM Works Up Close: Case Illustration	24
Lessons and Issues for Further Research	33
References	38
Appendix A	41
Tables and Figures	
Figure 1. SAM’s Logic Model: Leadership Development and School Reform	9
Table 1. Pipeline Outcomes of SAM’s Credentialing Program	14
Figure 2. SAM Student Outcomes	16
Figure 3. School Culture Trends in Mature SAM Schools	19
Figure 4. Inquiry Team Effects on School Culture	20
Figure 5. Inquiry Team Effect on Student Outcomes	21
Figure 6. Principal Effect on Inquiry Team Functioning	22
Figure 7. Facilitator Effects on Inquiry Team Progress	23
Figure 8. Marble Hill Teacher Reports on Assessment Use	28
Table 2. Scantron Results by Grade Level Equivalency in Reading	30
Table 3. Regents Performance	31
Figure 9. Marble Hill Student Outcomes	33

Leadership Development and School Reform through the Scaffolded Apprenticeship Model (SAM):

Introduction

Despite growing agreement among researchers and policy makers that evidence-based practice, collaborative learning, and distributed leadership are key to continuous school improvement, limited understanding exists about how schools and school systems can establish these conditions. Although the “terms” associated with these improvement strategies travel well, “the underlying conceptualization and thinking do not” (Fullan, 2005, p. 10). Nor do most studies of exemplary schools provide a theory of change for creating their effective cultures and practices.

The Scaffolded Apprenticeship Model (SAM) attempts to fill these gaps. It offers a theory of change and design for collaborative, evidence-based practice and broad leadership development that is beginning to show strong results. SAM integrates a university-based, degree-granting leadership development program with inquiry-based school reform. This marriage and principles for practice in each domain are grounded in lessons from research on administrator credentialing programs and on school reform. Through its credentialing program, SAM develops the capacity for a school leadership team, composed of teachers from different school units, to use data to identify student learning gaps and target interventions to expand the school’s “sphere of success.” SAM aims to develop school teams’ capacity to achieve significant improvements in student achievement and, at the same time, to develop a pipeline of school administrators equipped to lead inquiry-based reform in high-poverty urban schools.

This report evaluates SAM’s outcomes and describes how and with what challenges school teams develop effective inquiry practices and lead school reform. We first document the rationale for SAM in terms of key research findings that ground the program and describe its design and core principles for team practice. Then we summarize outcomes and experiences of SAM II participants and their schools, including a case study of one small high school that participated in SAM II and subsequent iterations of the program.¹ Our analysis spans a period of approximately four years, including the time SAM II teams participated in the credentialing program (January 2006-June 2007) and the subsequent two years. This time frame affords a look at the developmental trajectory of school leadership and change, as well as a fair short-term assessment of student and administrator pipeline outcomes.

The study’s findings have implications for ongoing SAM work in New York City and beyond² and for other initiatives that promote teachers’ use of student assessment data for continuous school improvement. Currently, two versions of SAM are being implemented in New

¹ SAM II is the second iteration of SAM that involved teams from 14 schools. Liz Gewirtzman and Nell Scharff of The Baruch College School of Public Affairs co-directed the work of five SAM II instructors working with these schools. SAM I was a pilot program with 4 schools led by Liz Gewirtzman and Nell Scharff of Baruch College and Ronald Chaluisan of New Visions for Public Schools.

² SAM is currently being implemented in Boston Public Schools through a partnership with the Boston Plan for Excellence (BPE) and in the San Francisco Bay Area through partnerships between the Bay Area Coalition for Equitable Schools (BayCES) and the Oakland, Berkeley, and Emeryville school districts.

York City schools. One is a certification version that replicates SAM II, successful completion of which results in an administrative credential through the School of Public Affairs, Baruch College.³ A non-certification version offers on-site facilitation of the inquiry team's work but not the structured assignments, intensive support, and accountability of the certification program. This study compares inquiry progress and outcomes for schools in their fourth year of the certification model with those with less than two years of the non-certification version.

Lessons from SAM are especially important in the NYC context where demand for administrators in high-need schools is expanding beyond pipeline capacity and where DOE policy mandates that all schools have an Inquiry Team charged with implementing the model of data-based decision making developed in SAM.⁴ Beyond NYC, demand for school leaders skilled in leading inquiry-based reform has grown through NCLB (2002) and state and local accountability systems that call for evidence of continual improvement in student achievement.

SAM's Rationale and Research Grounding

SAM puts forth a new paradigm for administrator preparation that involves school teams in using data-based inquiry to improve student achievement in their school. It addresses limitations of traditional credentialing programs for preparing administrators and takes on the challenges of developing evidence-based practice in schools.

*Limitations of traditional credential programs.*⁵ The critical role of leadership in organizational effectiveness is well documented within and outside education (Fullan, 2001; Harris, 2008; Senge, 1999). In particular, turning around troubled schools depends upon leaders who deeply understand the problems of change and know how to act strategically to build capacity for improvement. This entails building collaboration and the collective efficacy of a group to improve student learning (Sharratt & Fullan, 2005). Such leadership skills and stance are fundamental to improving education in high-need urban settings.

As knowledge of the nature of leadership for school improvement expands, it is clear that typical administrator credentialing programs are not well designed to prepare school change leaders, especially for inner city schools. Many programs have low admission and graduation standards, weak curriculum and instruction, and clinical experiences that are inadequate in quantity and quality (Levine, 2005). Not surprisingly, then, success in a university setting is a poor predictor of success in a school setting (Gladwell, 2008; Kane et al, 2006; Rockoff, Jacob,

³ SAM III (January 2008-December 2009) involved teams from 17 schools in New Visions' PSO (11) and in the ESO (7), including four schools that had been involved in SAM II. Nell Scharff of Baruch College trained six SAM III instructors to work with these schools. SAM IV was launched in September 2009. Further, during 2009-10, Brooklyn College and Lehman College are developing SAM as an option in their administrative credentialing programs.

⁴ In 2007-08, the New York City chancellor initiated the requirement that all schools create an Inquiry Team charged with using data to improve student outcomes. The policy is an attempt to scale up SAM; it replicates the design for composing teams and the inquiry model. This mandate is part of the Department of Education's *Children First Initiative*, which also increases principal authority and accountability for results and provides a range of diagnostic data to support school improvement.

⁵ The rationale presented here draws heavily on Gewirtzman's (2009) proposal to New York State for an Educational Leadership Program Enhancement grant to extend SAM to administrator credentialing programs beyond Baruch.

et al, 2008). Skills that enable success in typical administrator preparation programs do not necessarily correlate with those needed to lead schools, particularly high need ones. The intensity and quality of a practical experience component or internship vary widely across programs, but they cannot replicate the hands on experience of leading a school change process.

SAM is designed to address the gap between individual learning in a credentialing program and conditions of leading school improvement. It does this in several ways. First, it involves a team that includes promising teacher leaders across the school. Research shows that “distributed leadership” is key to continuous and sustained improvement (Spillane, 2006) and that, especially in high schools, school reform efforts can be stymied by the segregated worlds of subject areas (Siskin, 1994). SAM prompts schools to put together teams of aspiring leaders who span school units – subject departments or Small Learning Communities (SLCs) in secondary schools and grade levels in elementary schools. This design establishes conditions for the team’s leadership to span the school, as well as for participants to learn how to function in a collaborative leadership team.

Second, SAM’s curriculum is designed around real problems of improving student achievement in schools, established standards for team functioning and leadership practices, and assignments focused on the team’s school and student population.

Ultimately, SAM develops new school leaders’ capacity to work effectively in school teams that use evidence to continually improve student achievement. Although SAM intends that some graduates will become principals or APs in their own or another school within a few years, the program also expects that some participants will continue as strong leaders of inquiry work in their school.

Challenges of inquiry-based school reform. Despite policy demands for schools to use assessment data to evaluate and improve performance (NCLB, 2002) and local education leaders’ enthusiasm for this theory of change (Archer, 2005), research on schools’ use of evidence to make instructional improvements suggests that few do so. Challenges stem from incoherence between administrators’ and teachers’ conceptions of useful data, difficulty translating knowledge of student learning gaps into instructional interventions, and teaching cultures and school politics that maintain the status quo (Lachat and Smith, 2005).

Research finds that, for one, educators and administrators tend to hold different conceptions of what constitutes valuable evidence (Coburn and Talbert, 2006). As a result, district data systems and designs for schools’ use of these for instructional improvement often are out of sync with teachers’ needs. Second, teachers typically have little experience or support in using assessment data to detect specific student learning gaps and to design or identify effective instructional interventions (McLaughlin and Mitra, 2003; McLaughlin and Talbert, 2006), and tend to resist evidence use as a means for improving instruction (Ingram, Louis, and Schroeder, 2004; Supovitz and Klein, 2003). Third, when pressed to teach to their state’s standards for content instruction and to follow district pacing guides, teachers often feel they have little slack to diagnose and address the learning needs of students who fall far below grade-level preparation.

In addition, norms of privacy in teaching work against educators developing collective responsibility for improving instruction (Little, 1982). Current federal policy under NCLB that defines “teaching quality” as individual and based in formal education does nothing to challenge teacher autonomy norms; while proposed state and local merit pay schemes that would isolate individual teachers’ value to student learning are likely to further inhibit teacher collaboration and shared responsibility to improve student achievement. Challenges of developing teachers’ capacity and desire for evidence-based school improvement are particularly daunting in high schools due to their size and organizational complexity, subject-specific assessments, sub-cultures that resist school-wide instructional interventions, and typical teacher-tracking practices (McLaughlin and Talbert, 2001; 2007; Talbert, 2002).

School leaders who take on the challenge of developing evidence-based practice in education thus encounter a wide range of technical, organizational, cultural, political, and policy obstacles (Talbert and Wood, 2007). The SAM program partners with schools to help develop their leadership capacity for school culture change. As elaborated below, the administrator credentialing program is designed both to address challenges to evidence-based practice within participating school teams and to prepare the teams to lead change in their or another school.

SAM’s Theory of School Change and Design

SAM’s stance on the problem of change is that every school has a “sphere of success” – a group of students with whom the school is currently successful – and the challenge is to continually expand the sphere by using evidence of struggling students’ skill gaps and addressing them. This agenda for school reform seems fairly straightforward. However, because its objective challenges current practices and thinking about quality education, as outlined above, it requires a program that strategically shifts teachers’ and administrators’ thinking about why students struggle to succeed and how teachers and schools can address their needs.

SAM’s design features teams, tasks, and tools to develop leaders capable of moving the culture of a school toward conditions for continuous improvement. Each is designed to both challenge participants’ habits of mind that maintain a sphere of success in the school and to develop their skills in leading colleagues toward new perspectives and practices. SAM also includes a facilitator to guide the work and keep the team centered on student learning.

The Inquiry Team. As a model for developing school leadership capable of addressing school failures, SAM creates an “inquiry team” comprised of teacher leaders representing a broad array of school units. The inquiry team is a vehicle for distributing school leadership broadly. At the same time, it creates a community of practice around the work of improving student learning through data-based inquiry. The SAM curriculum supports the inquiry process and also guides the team to think systemically about the problem of school change and leading learning within a school. SAM participants learn how to work as a leadership team 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 design for developing teachers' collaboration to improve student achievement, SAM focuses the team's joint work around specific tasks. Teams are prompted to:

- Identify gaps in skills of targeted students outside the sphere of success and gaps in instruction to address their learning needs;
- Design high-leverage instructional and programmatic responses that close the skill gaps and accelerate student achievement; and
- Engage colleagues in inquiry.

These tasks anchor SAM's curriculum and assignments in three phases: research, action, and leadership. In the research phase, the team studies their school through the lens of a specific skill gap for struggling students, coming to understand how patterns in decision-making school-wide reliably produce the current sphere of student success. In the action phase, the team learns from iterative phases of action research to improve outcomes for target students and to improve one or more decision-making systems schoolwide that have produced underperformance for targeted students in the first place. In phase 3, the team focuses explicitly on leading colleagues to conduct and own the inquiry process, so that evidence-based improvement can be sustained. 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 the skill gaps of a small group of struggling students prompts participants to shift their thinking from teaching to learning and sets the stage for their purposeful 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 evidence-based improvement.⁶ A SAM team (small strategic group in their school) is 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.⁷ 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. For one, it shifts their focus from assumptions about reasons for student failure to evidence-based knowledge around specific skill gaps. Second, by investigating where and how a student can learn a specific skill in the school, their view of instruction shifts from curriculum delivery and teacher expertise in a subject to student learning.

Tools to frame and support SAM teams' work. SAM'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* they can change these conditions.

⁶ See Talbert and Scharff (2008).

⁷ 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.

Key tools include:

1. 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.
2. Readings aligned with SAM modules.
3. Protocols for coaching and inter-visitations.

Low inference transcripts are used to help shift participants' views from teaching to learning and to help them to disentangle assumption from fact (a pre-requisite for opening their minds to alternative ways of seeing target students and their potential impact on those students). LITs also provide valuable data for understanding the curriculum as taught and experienced by students – rather than as it exists on a map or in participants' minds. When SAM participants analyze transcripts through the very specific lens of what target students do not know to see if it is taught to them, they are directly confronted with the reality that students typically do not have opportunity in any of their classes to learn the skills they lack. In this way, this tool supports the core task of understanding how current conditions produce current outcomes.

Readings are designed to support a changed view or new idea and are usually scaffolded by an experience, rather than vice versa, as is usually the case with other credentialing programs. Carefully selected readings for each SAM module help participants to make cognitive shifts essential to the work. For example, during the research phase, participants 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 Lusk, 2002).

Protocols for facilitator coaching with team members and inter-visitations between schools guide facilitators' action to support individual and school cultural shifts. Individual leadership coaching by a trained SAM facilitator provides participants an opportunity to work on increasing strategies for managing what is most difficult for them personally in the work of doing and leading school improvement. The coaching protocol leads the participant 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 through both getting on-site feedback from the SAM colleagues from other teams and developing a lens and norms for providing useful feedback to colleagues. The protocol establishes a structure and norms for a productive visit: the visited school articulates a problem of change and invites a team or teams to participate in the problem-solving process. The visiting team collects data and offers it to the hosts, who can then utilize this to publicly work towards solving school-wide problems. The SAM practice and protocol are designed to support a shift toward public learning and the development of leaders' skills in giving honest and actionable feedback.

Facilitators. A SAM facilitator supports the team's use of data to design and monitor instructional interventions to close skill gaps, to improve decision-making systems that led to the

identified gaps, and to lead colleagues to do the same. The facilitator keeps the team on task, provides feedback on the quality of their work, and pushes them to develop the discipline of inquiry. For example, holding the team accountable for rigorously assessing interventions according to evidence of accelerated student learning in one measurable skill pushes participants to examine current practices and the decision-making systems that underlie these practices. In turn, the work provides a leverage point for individual and organizational change.

Facilitators are expected to interpret and help to enact the SAM program with each participating school team. Because coursework is specific to each team and its student achievement gaps, facilitators need to be deeply grounded in core program principles and clear on how the tasks and tools both leverage and support learning in the teams. Toward this end, the SAM program has built in a day per week in which facilitators convene to develop seminar lesson designs, review team work on assignments and calibrate standards for assessments. These practices use ensure program quality. Facilitators also raise issues from their work with school teams and administrators that focus discussion of strategic responses, critical to their success in supporting the change process. A facilitator community of practice is a key resource for the SAM program, enabling it to make ongoing refinements that advance the work and to support facilitators' ongoing learning to improve their practice with schools.

Figure 1 presents the overall design and logic model for SAM II and a schema for the evaluation. It shows partners and their responsibilities in developing and implementing SAM over time and specifies assumptions about how the model is implemented in schools and with what intermediate and ultimate outcomes. In this visual representation of SAM's logic model, "SAM Inquiry Model" refers to the program features just described.

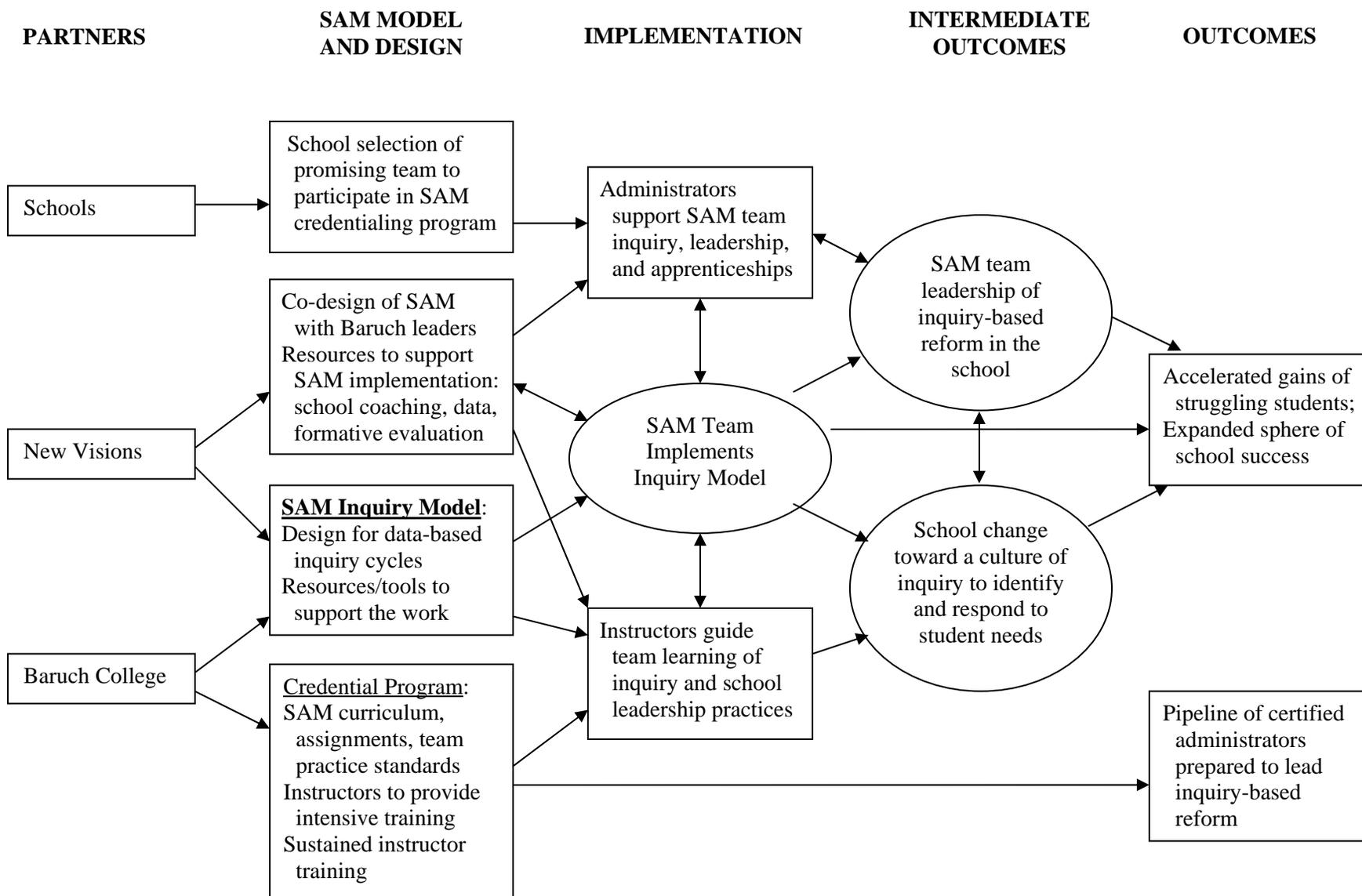
The evaluation is designed to assess the model's hypothesized cause-effect relationships over time, as well as to provide ongoing feedback to SAM and NV leaders on their efforts to support the development of evidence-based school cultures. This report summarizes results to date of our assessment of SAM outcomes for prospective administrators and for students in participating schools, intermediate outcomes of school leadership and culture change, and conditions that affect teams' progress on implementing SAM.

The report addresses three broad questions:

- To what extent did individuals and schools participating in SAM II realize the intended outcomes of attaining administrative credentials and leadership positions and improving the school's sphere of student success?
- Does a school's progress toward an inquiry culture and SAM team leadership development make a difference for student outcomes?
- To what extent and how does the principal and team facilitator make a difference for a school team's progress on inquiry and leading school change?

We take up each of these questions in turn, as separate sections of the report. A case study then illustrates how the model has worked to developed inquiry leadership and evidence-based practice in a school that has participated in SAM over the past four years.

Figure 1. SAM's Logic Model: Leadership Development and School Reform



Research Methods, Data, and Analyses

We evaluate SAM's theory of action – and the particular questions outlined above – using evidence from two successive iterations of SAM.⁸ We draw upon data from the fourteen high schools that participated in SAM II,⁹ especially four of the schools that extended their participation into SAM III, and from our ongoing evaluation of SAM III and the work of Inquiry Teams across schools in the New Visions Partnership Support Organization (PSO)¹⁰. Quantitative and qualitative data collected during four years provide breadth and depth of analysis of SAM outcomes, team leadership of school culture change, and conditions that made a difference for team progress in implementing the SAM model.

Quantitative analyses draw upon record and survey data developed through the SAM II and SAM III evaluations, including:

- SAM II participant pipeline outcomes two years after program completion
- Students' "On Track" statuses for all New Visions high schools (Winter 2008)
- Teacher survey data for SAM II-III case study schools (annually, May 2006-09)
- SAM II team survey data (February 2007)
- Inquiry Team survey data for all New Visions schools (May 2008, 2009)¹¹

Qualitative data on SAM teams' experiences implementing the inquiry model and leading school change come from:

- Focus groups with 9 of the 14 SAM II teams (Spring 2007)
- Annual principal interviews in five SAM II former-Region 1 schools (2005-06, 2006-07, 2007-08)
- Repeated principal interviews in 10 SAM III case study high schools (twice annually, 2007-08 and 2008-09)
- Focus groups and interviews with team members in 10 SAM III case study high schools (2008-09)
- Longitudinal case studies of four schools involved in SAM II and SAM III (2006-09)

Samples and types of data used in this report vary according to the purpose of analysis. Here we provide a "roadmap" of data used to address each evaluation question and a brief description of analysis techniques.

⁸ We focus on outcomes of SAM II (January 2006-Fall 2008) and leading indicators of progress in SAM III (January 2008-December 2009). SAM's first iteration (SAM I: Summer 2004-2005) involved four schools in a pilot program. SAM IV began in Fall 2009. The evaluation continues to follow SAM III and is beginning to document the work of SAM IV teams.

⁹ The SAM II schools included two large restructured high schools (one in Queens and one in Staten Island), five small high schools in former Region 1 of the Bronx, and seven small high schools in the Autonomy Zone (renamed Empowerment Schools in Fall 2007). Each of the large schools had its own SAM instructor and on-site seminars; the five Region 1 schools formed a cohort and shared an instructor, and the seven Empowerment Schools formed another cohort with two instructors.

¹⁰ In 2007-08, through an application process, New Visions became one of several private PSOs, one of three forms of School Support Organizations (SSOs) created by the DOE's restructuring of NYC school governance. New Visions currently serves approximately 75 schools that opted into the PSO.

¹¹ We use available data for the 63 schools that were in the PSO during both 2007-08 and 2008-09.

Examining evidence of SAM's effectiveness. We assess administrator pipeline outcomes using New Visions' record data for all SAM II participants, including all individuals who began the program in each of the fourteen participating schools. Data for each participant include: a) whether or not s/he completed the program and graduated; and b) her/his position – teacher, AP, or Principal – as of fall 2009. We followed up with SAM II facilitators to determine where those in administrative positions were currently located and, for the large high school that participated in the certification program, to identify who had become Director of a School Learning Community (SLC).¹² We assess pipeline outcomes in terms of the program's credentialing and administrator placement rates.

In assessing *student outcomes*, we use New Visions' database for individual students in each of the PSO schools. Results reported here use each student's 8th grade score on the state ELA tests and "On Track" status for his/her graduating cohort as of winter, 2008. New Visions' On Track ratings use grade-level criteria for a student's course completion and Regents scores to classify the student as off track for graduation, almost on track for graduation, on track for graduation, and on track for college readiness. Importantly, this metric was *not* designed for the purpose of evaluating school outcomes. Rather, New Visions provides schools with data on the distribution of individual students' status by grade level to help guide their decisions about how to get students on track for college readiness. For example, a school may begin to offer Regents exams in the 9th grade after seeing their data.¹³

Our analysis compares On Track statuses for New Visions schools that had participated in SAM for four years with those for schools with no SAM experience.¹⁴ In order to control for school differences in student achievement when they entered the school, we examine student outcomes for students who entered the school with 8th grade ELA test scores below Proficient (Basic or Below Basic levels combined).¹⁵ We examine these students' patterns of performance across grade levels – on average for the two groups of schools – testing the hypothesis that the veteran SAM schools will have significantly greater proportions of 11th graders on track to graduate. Increasing proportions of students on track across cohorts would suggest a positive school effect for students who were struggling academically when they entered the school.¹⁶

¹² An SLC functions as an autonomous program within a large restructured high school; it has a curricular theme and a devoted faculty across core academic subjects, and it serves 300-400 students in grades 9-12. [Note: one of the two large restructured SAM II high schools elected not to have teacher candidates in the certification program and therefore is excluded from this analysis. Its SAM II teams included credentialed APs who were charged with leading the school's transition from a comprehensive high school to themed SLCs or "houses". The school embraced SAM's theory of action, followed its curriculum, and had intensive facilitator support during SAM II and beyond. Currently, each of the SLCs in the school constitutes an inquiry team charged with improving their students' success.

¹³ Note that the On Track data are reported for student cohorts. In summaries here, we label the cohorts according to students' expected grade as of Winter 2008; for example the 2009 cohort is labeled "11th graders." Nevertheless, all students in a cohort are included in the data for their expected grade level; students not promoted would appear as "off track."

¹⁴ The comparison schools include 62 "non SAM-cert" schools within the New Visions PSO in 2008-09 (excluding eleven schools that were participating in the SAM III certification program).

¹⁵ Close examination of differences between the two groups of schools reveals that, on average, the SAM schools had higher proportions of Below Basic students and significantly higher proportions of Below Proficient students in the 11th grade.

¹⁶ This method of "synthetic cohort analysis" examines change in student outcomes across graduating classes in order to infer a school effect on student performance as they move through a school. It assumes that the school's

Note that this analysis strategy cannot take into account a school's innovations directed at entering student cohorts and thus will underestimate school effects other than those that show up across the cohorts. This problem applies equally to the two groups of schools, however. Ideally, an evaluation of school effects would examine the progress of individual students over their years in the school – a strategy that requires at least three years of individual student data for each school and will be possible within the next year or so.

Assessing school culture and leadership outcomes. Our evaluation of *school culture change* through SAM uses data from schools that have been involved in the program for at least three years. Included are two large restructured high schools and two former Region 1 small schools that have continued with the SAM III certification program. Trend data are school means on a teacher survey measure of their use of student assessment to inform instruction (see Appendix A for items that make up our measure of “Culture of Assessment Use”). Qualitative data from interviews and focus groups with school teams in the Region 1 SAM II cohort complement results of the survey analysis by illustrating how the teams experienced the program and what they took as evidence of change.

Assessing principal and facilitator effects on inquiry progress. We use the broader NV-PSO sample and Inquiry Team (IT) survey data for 2007-08 and 2008-09 to estimate effects of principal support and facilitator support on IT progress. We use the IT members' ratings of their team functioning and inquiry leadership in the school, as well as their ratings of kinds and extent of support from their principal and from their New Visions facilitator. Included are schools for which we have two years of data for at least two IT teacher members (N=38); school scores on survey items are mean IT member responses. Our analysis tests hypotheses in SAM's theory of change (Figure 1). The model predicts, for example, that a facilitator's support of the IT's focus on results and data use predict team leadership outcomes. In estimating facilitator effects on team performance in 2009, we control for prior (2008) measures of team performance. Structural equation modeling (SEM) techniques are used to estimate effects between variables measured by multiple survey items (see Appendix A for items used to specify each variable.).

Case illustration of how SAM works up close. This case study school was part of the SAM II Region 1 cohort and has brought new teams into the SAM IIIA and SAM IIIA credentialing programs. As part of our evaluation case study sample, the school has provided annual teacher surveys and interviews with administrators and SAM participants over four years (Spring 2006 –Spring 2009). We draw upon all of these data, as well as interviews with SAM facilitators, to capture the school's progress through SAM. Our case description draws upon the SAM II and SAM III teams' reports on their inquiry work.

Evidence of SAM's Effectiveness

SAM's goals of developing both collaborative school leadership for inquiry-based reform and a pipeline of administrators for NYC schools are somewhat at odds. As SAM teams develop commitment and skills in leading change in their schools, individual team members may opt not to pursue an administrative position in another school. SAM graduates' opportunities to take on

effect results in increasing success among students in successive cohorts. As the New Visions database matures, we will have three years of individual-level data needed to assess cohort-specific trends at the individual student level.

an administrative position in their own school are inherently limited, so pipeline outcomes may extend longer into the future than is typical of administrator credentialing programs. Nonetheless, we expect to see strong positive outcomes for students of SAM teams' efforts over a period of 3-4 years.

SAM as an administrator credentialing program: pipeline outcomes. SAM participants' short-term career outcomes reveal that their commitment to leading inquiry-based reform in their school is generally greater than the impetus to take an administrative job in another school.

Despite extremely high rates of certification across SAM II cohorts (96 percent overall), only about a third (36 percent) of teachers certified were in administrative positions two years later (see Table 1). Among those who had become administrators, most were in AP positions in the same school; one became principal in the same school. Only 5 of 50 SAM-certified educators took administrative positions outside of their school: 4 from former Region 1 schools and 1 from the large high school. One opened a new school as its principal, two became principals of existing schools, and two became APs in other schools.

Pipeline outcomes are stronger if we count the 9 SAM participants in the large restructured high school who became Director or Co-director of their SLC. Although an SLC directorship is not formally an administrative position, it carries considerable responsibility and resource control. Indeed, the SLC Directors' leadership was pivotal to the school's successful transition from a traditional department structure to small themed learning communities serving around 400 students each. The principal regarded SAM as a key vehicle for developing the leadership skills and legitimacy of the teachers who became SLC Directors. Ultimately SAM graduates in positions as Directors of interdisciplinary SLCs led their colleagues in using data-based inquiry to improve instruction for students in their SLC, with support of APs' content instruction expertise and supervision.

Even using the more liberal criterion of including SLC Directors in the large high school, just over half (54%) of the teachers certified through SAM II became administrators or formal leaders in their own or another school within two years. This rather low placement rate is unsurprising, considering that teams included considerable numbers of individuals who were not interested in pursuing administrator positions in the near future. A spring 2007 survey asked SAM participants to indicate their interest in pursuing a position as AP or Principal in the future. On a scale of 1 (no plans to pursue) to 4 (definitely will pursue), just 21 percent gave a 4 rating for AP and 17 percent gave a 4 rating for Principal. Conversely, 18 percent rated AP as 1 and 30 percent rated Principal as 1 – quite high proportions of teachers certified through SAM without interest in becoming an administrator. In response to a survey question of *where* they would prefer taking a future leadership position, 61 percent of all participants indicated strongest preference for their own school.

Participants who were not interested in pursuing administrative positions stated two main reasons for joining SAM. Most wanted to help lead instructional improvements in their school and joined SAM in order to develop their leadership skills. Their principals had tapped them as budding teacher leaders. Among them were teachers who had fewer than five years of teaching experience and regarded the SAM program a cost-effective path to certification should they

choose to pursue administration in the future. Their horizon extended as far as participating in a school leadership team that would collaborate to improve teaching and learning in their school.

Table 1. Pipeline Outcomes of SAM’s Credentialing Program: Participation, Certification, and Placement by SAM II Cohort

	Region 1 (5 small high schools)	Empowerment Schools (7 small high schools)	Large restructured high school (9 SLCs; 4 SAM teams)	TOTAL
Participants¹⁷	22 teachers	15 teachers	15 teachers (7 APs)	52 teachers
Certification (completed by summer 2008)	22 (100%)	14 (93%)	14 (93%)	50 (96%)
Placement in administrative position (percent is of all certified)	3 Principal; 5 AP (32 % total)	5 AP (33%)	1 Principal 4 AP (21%)	18 Principal or AP (36%)
[SLC director]			[9 SLC director (64%)]	[27 including SLC directors: (54%)]

Regardless of their immediate aspiration to an administrative position, most SAM participants have thought that the credentialing program offered credibility to their team’s leadership, prepared them to lead school change, and developed their identity and commitment as a school leader. A teacher in a SAM II Empowerment School put it this way:

I think that the administrative credential part [of SAM] really has sort of lent it a gravitas and created a lot more buy-in for us in terms of taking on an insane amount of work and dedicating a lot of time in a way that an administrator does to a school, and sort of an ownership for the way the school is running in the way that an administrator does. And I think that while teachers are definitely interested in developing that, in developing the school and in helping kids, one of the most valuable things about this program has been its leadership development for me. And I’m not sure what the program would look like without that aspect [SAM without its credentialing program]. And I’m not sure it would have moved me forward as much if it hadn’t had that aspect also.

¹⁷ Numbers are for SAM participants in the credentialing program. (Note: In some schools, credentialed administrators also sat in on seminars, e.g., seven APs in the large high school attended seminars and did some of the assignments.)

In this view, SAM’s credentialing components and its approach to school reform are synergistic: together they set conditions for developing school leadership capable of improving student outcomes.

The long-term pipeline outcomes of SAM II are likely to be considerably stronger than those documented in Table 1. As young participants gain experience in teaching and leading inquiry with colleagues, their career aspirations may well turn to administrative leadership. SAM graduates will be well equipped to move into administrative positions that become vacant in their schools. In the short term, we find evidence that SAM’s design for collaborative leadership of inquiry-based reform is paying off for students – the ultimate goal of any leadership development program.

SAM as a school reform model: improved student achievement. SAM’s theory of action posits that school teams using its model and curriculum for inquiry practice and leadership development will gradually expand their school’s sphere of student success. Using data described earlier, we examine grade-level outcomes for students who entered the school with less than proficient ELA skills. We expect to see increasing proportions of students on track from 9th to 10th to 11th grades.¹⁸ Comparisons of cross-grade trends for schools that had participated in SAM for three years with those of similar schools that did not participate in the program assess a SAM “effect” upon student achievement.¹⁹

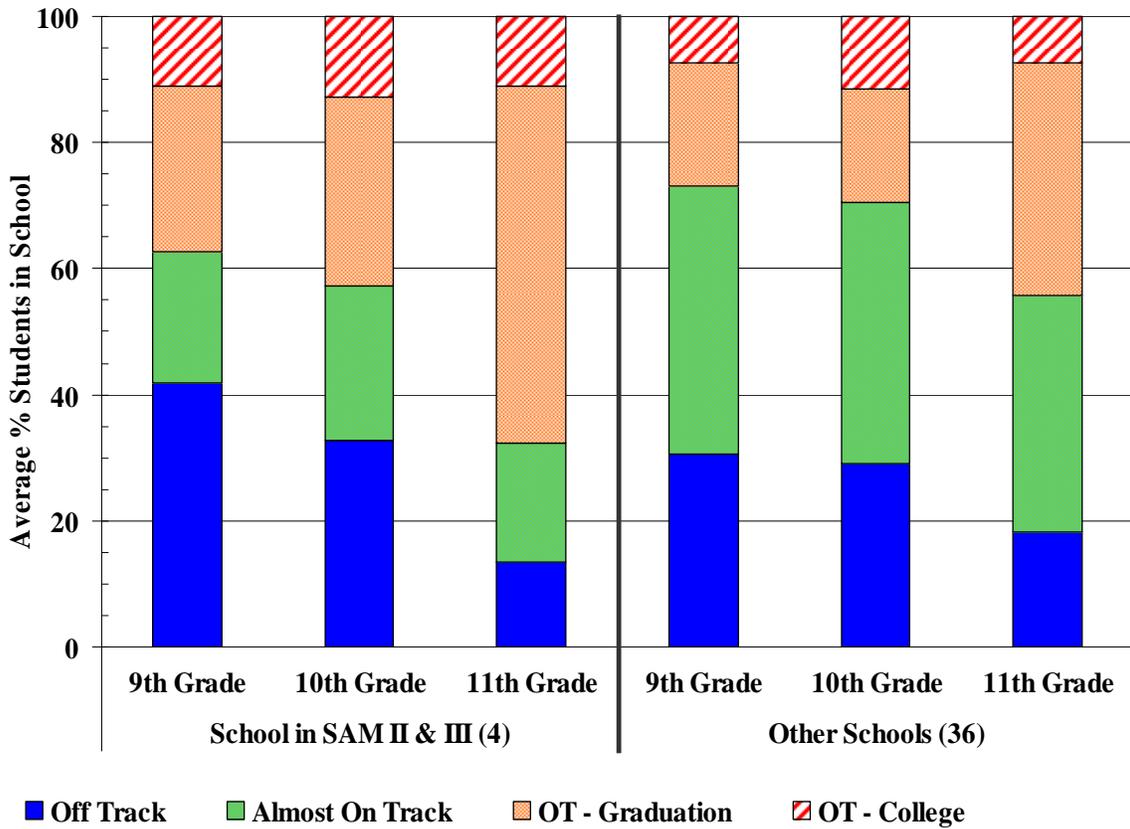
Data summarized in Figure 2 support the hypothesis that SAM schools better promote success among students who enter at risk of not graduating high school. It appears that the four schools with sustained SAM participation far exceed the typical school in bringing these students on track. On average, students “off track” to graduation declines from 42 percent among 9th graders to 13 percent among 11th graders. The percent of students on track jumps from 37 percent to 68 percent between 9th and 11th grades. This trend cannot be explained by a difference in percent of all students in each cohort who scored below proficient before entering the schools. In fact, the proportion of current students that scored below proficient in 8th grade was highest for the 11th grade cohort (70 percent, compared to 62 percent for the 10th grade cohort and 55 percent for the 9th grade cohort).

The positive trend for SAM schools is significantly greater than for comparison schools. Higher proportions of 9th graders were on track at the end of their first year (37 percent versus 27 percent in non-SAM schools), and the increase in percent on track between 9th and 11th grade was substantially larger (31 percent for SAM schools versus 17 percent for non-SAM schools). Comparison schools had similar proportions of students in each cohort who entered below grade level, though for 11th graders the percent is lower than for SAM schools (52 versus 70 percent). The direction of difference rules out the possibility that SAM schools had better outcomes because they had smaller proportions of students who were struggling academically in this cohort or that they had higher dropout rates for struggling students in the cohort.

¹⁸ As noted earlier, our data set does not yet have 3 years of longitudinal data, so we use this approach (synthetic cohort analysis) to estimate trends for students who entered the school with 8th grade performance levels below Proficient.

¹⁹ For each group of schools we examine the percent of students at each grade level that scored below proficient in 8th grade to ensure that differences in percent on track are not due to low-performing students dropping out.

Figure 2. SAM Student Outcomes: “On Track” Performance for Students Below Proficient Level in ELA in 8th Grade (Before Entering the School), by Student Cohort



SAM teams promoted target students’ academic growth and improved high-stakes test performance through a wide range of instructional responses. Many teams also engaged students in analyzing their data and identifying learning targets to address gaps in their test performance. These team members reported seeing a shift in students’ motivation to master course content. Some commented on changes they had seen in target students’ engagement and identity as learners. A teacher in a small Bronx high school used an example of two girls sharing their scores on a diagnostic ELA assessment:

And one received a high DRA [Developmental Reading Assessment]. And the other one asked – and she wouldn’t have asked this before – “How did you do that? Because *I* want to do that. How did you get that DRA to be so high?” So what I found happening with the kids was that they started looking at themselves differently... That was the biggest piece. Because once they start doing *that*, then we have them.

As SAM teams involved target students in analyzing their own performance on various assessments, the students developed ownership of their learning and a new understanding of the consequences of their performance on high-stakes assessments.

SAM's theory of action attributes schools' success on student outcomes to the teams' work to: a) identify and address target students' skill gaps and move their performance; 2) use this knowledge to make changes in curriculum, student placement, or other structural or programmatic elements that improve the system for other students; and 3) lead their colleagues to implement system changes and develop inquiry practices. We use available data to examine school trends on leadership development and culture change in these directions and to evaluate context effects on teams' progress.

School Culture and Leadership Outcomes

The actual work of SAM – what comes of the inquiry process and how the team leads school change – is unique to a school, since each has its own student population and systems that maintain a sphere of success. SAM prompts and supports a team to use its school's data to identify which, why, and how students struggle and to develop effective instructional and programmatic responses.

Most of the fourteen SAM II school teams identified skill gaps in reading or writing – gaps that hindered the target students' performance in all content areas and on Regents exams. Exceptions were teams within the two large high schools: 5 of 8 SLC teams in one school and 2 of 8 teams in the other focused on a math skill gap. Target students in these SLCs were weakest in mathematics, in part because they had chosen a curricular theme that fit their interests and skills; for example, in one school's Academy of Fine and Dramatic Arts students struggled less in language arts than in math.

The teams identified a range of school “system” conditions that inhibited target students' skill development, including curriculum gaps, teacher assignment patterns that disadvantaged struggling students, and inadequate creation and flow of student assessment information. They designed structures and policies to address such problems and gradually involved colleagues in using inquiry to identify and respond to student skill gaps. In small schools, SAM teams reached out to colleagues who had target students in their classes and also presented data on their inquiry work to the whole faculty. In large schools, teams involved their SLC colleagues in interventions to address target students' skill gaps and in reviewing assessment data on learning outcomes; at the end of the year each SLC team presented results of their inquiry work to faculty in the other SLCs. Both large and small schools that persisted with SAM brought successive cohorts of teachers into the credentialing program, developing a critical mass of inquiry leadership.²⁰

Schools' participation in SAM appears to have changed their professional culture in ways predicted by its theory of action. Evidence includes trends toward evidence-based practice in schools with sustained SAM participation and statistical effects of team functioning on school outcomes within a two-year time frame.

SAM schools moved toward a culture of inquiry. Schools with sustained SAM involvement over nearly four years have moved steadily toward a culture of inquiry-based improvement. Teacher survey data for these four schools show incremental growth on our

²⁰ See under separate cover case studies of SAM teams' work and consequent school change in a large high school involved in SAM since 2004 and in a small high school involved since 2006.

measure “culture of assessment use.” Figure 3 shows results for two large high schools in Staten Island and Queens, each with multiple inquiry teams working across their houses/SLCs (Schools A and B) and for two small high schools in the Bronx (Schools C and D, formerly in Region 1 and now in New Visions’ PSO).

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). As is common in many schools, administrators reviewed scores from standardized tests – in particular Regents examinations – and received a year-end “scorecard” for the school and for particular content areas. Individual teachers and subject departments had considerable latitude in deciding whether and how to use finer-grained interim assessments. In these three schools, teachers were almost as likely to disagree as to agree with statements that assessments were used to inform instruction. Since the fourth long-term SAM participant (School D) had a tradition of assessing individual student performance through portfolios submitted twice a year, teachers’ initial ratings of assessment use were relatively high. Through SAM, however, the school made a qualitative shift toward using fine-grained skill assessments to identify and hone in on learning targets for struggling students.

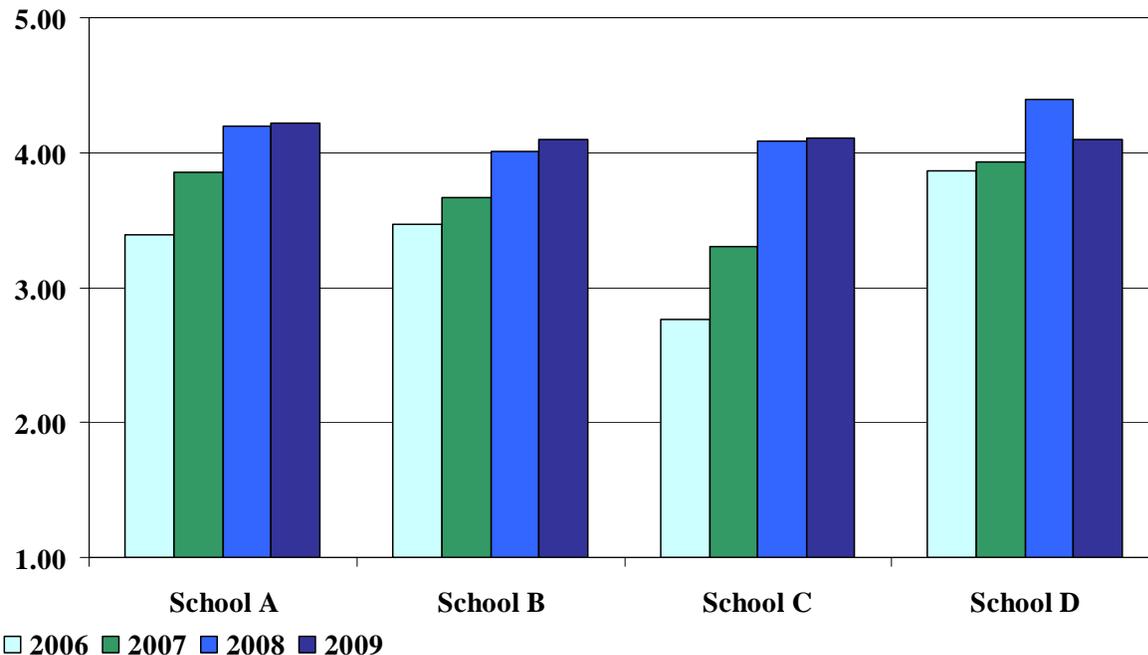
By their third year of SAM, teams in each school had involved colleagues in using SAM’s inquiry model, as reflected in teachers’ school-wide 2008 ratings on the survey scale. Across the schools teachers had learned to: a) identify target students who were not succeeding despite regular attendance; b) assess skill gaps; c) explore extant opportunities to learn needed skills in the curriculum and in class; d) develop strategic instructional responses; and e) evaluate learning outcomes of the interventions. Depending on their intervention’s success, teacher groups would either share evidence with colleagues or refine the intervention for all or some of the target students until the students had mastered the skill. Much of this work focused on student skill gaps in literacy, such as identifying the main idea in a text or knowing academic vocabulary in or across a subject area.

Data from our spring 2009 teacher survey in veteran SAM schools suggest that culture change plateaus after three years and, significantly, that the new inquiry norms and practices are being sustained. School means on the Culture of Assessment Use survey scale were around “4” in 2008 and 2009, indicating that teachers overwhelmingly agreed with the statements about assessment use in the school. Of course, the survey measure does not capture ongoing refinements in assessment use within the schools or the deepening of inquiry work.

Survey trends lend support to SAM’s theory of action, demonstrating the expected intermediate outcome of a developing inquiry culture in schools. As elaborated below, case studies of these schools and recent survey data for Inquiry Teams in a broader sample of New Visions schools provide further evidence that culture changes came about through SAM teams’ leadership.

SAM teams led inquiry-based reform in the school. Development of an inquiry culture depends upon the SAM team’s use of data to increase student success, sharing of evidence with colleagues on high-leverage interventions, and success in leading colleagues to adopt inquiry practices. Teams in each SAM school led colleagues to transform how they think about academic weaknesses and how they use assessment data to move struggling students.

**Figure 3. School Culture Trends in Mature SAM Schools:
Teacher Reports on Assessment Use**



Note: bars represent mean teacher responses to two 5-point Likert scale survey items that make up the “culture of assessment use” scale: a) “we use a variety of assessment strategies to measure student progress;” and b) “this school uses assessment data to evaluate teachers’ instructional practices.”

In describing how the team was challenging their colleagues’ beliefs, a SAM II participant in an Empowerment School commented:

And I think everyone on staff is at very different points on the question of “are students failing because they’re lazy?” That’s something we’re trying to move them on. But I feel like there’re opportunities to kind of give real-life stories like “after pulling out this child so many times...when I asked the whole class to make predictions, instead of sitting there and doing nothing he picked up his pencil and wrote something down.”

SAM teams worked to develop their staff’s appetite for using assessment data to evaluate interventions. A teacher in a Bronx SAM II school described in a focus group how her team had worked to develop ownership of inquiry among colleagues:

Initially we didn’t have a lot of buy-in. It was “just another team doing something else” in this school, and how did *we* “get selected to do it?” – and all those issues that come up when people are selected to do a particular program in a school. However, we *knew* that, so we took steps in this direction to have people buy into what we were proposing. And one of the things that our team did was to convey to the staff that these results and this accomplishment [moving the target students] was “not because of *us* but because of *you*

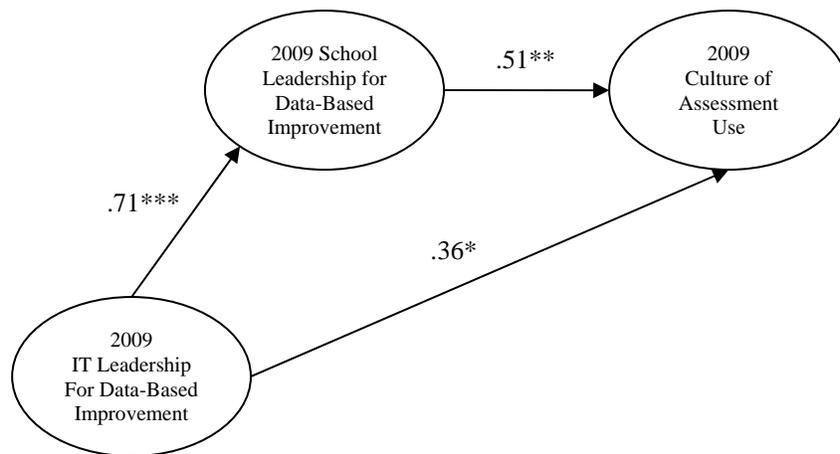
guys.” So that once we started talking about this as a community effort and not just five people doing all of this, I saw the shift and buy in. And when we presented the last time, I could *see* it. I could see the body language, that people were just very interested and wanting to know the numbers...At the beginning it was like “Okay, we’ve got to now listen to these folks with their data.” Now it was more about “Yeah, I want to really see what’s going on!” There was a shift in the attention.

Through team presentations of data to school staffs and “low-inference transcripts” of target students’ classes, SAM teachers focused their colleagues’ attention on students who in the past would have fallen through the cracks.

Data from our 2008 and 2009 Inquiry Team Survey provide additional evidence that the SAM teams are leading change toward an inquiry culture in their school. The SAM teams’ self-ratings on survey items that make up a “Leadership for data-based improvement” scale are significantly higher than self-ratings of Inquiry Teams not participating in the program. In 2008, when the SAM schools were in their third year of the program and non-cert New Visions schools were just beginning to engage with the model, the SAM teams rated themselves 4.4 on the 5-point scale, compared to 4.1 on average for non-SAM teams. This difference increased from 0.3 in 2008 to 0.4 in 2009.

Survey data also support the hypothesis that a team’s leadership is instrumental in bringing about school culture change. Inquiry Team self-ratings on the “Leadership for Data-based Improvement” scale predict scores on the “School Leadership for Data-based Improvement” and “Culture of Assessment Use” scales. Coefficients are statistically significant (see Figure 4). These results show that the *extent* to which a team takes an inquiry leadership role in their school predicts their school’s outcomes of school culture change, as measured by these leading indicators.

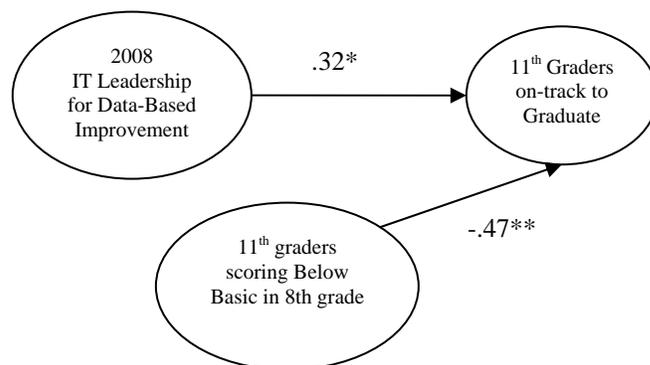
Figure 4. Inquiry Team Effects on School Culture (N=38 teams)



Further, teams’ self ratings on the inquiry leadership survey measure predict 11th grade students’ on-track outcomes, with controls for percent students with Below Basic performance in 8th grade (see Figure 5). These data suggest that a team’s inquiry leadership expands the school’s

sphere of student success. Schools with teams scoring high on inquiry leadership have greater proportions of students on track to graduate by 11th grade, after school differences in the students' performance prior to entering the school are taken into account.

Figure 5. Inquiry Team Effect on Student Outcomes: Percent 11th Graders on Track with Controls for Percent Scoring Below Basic as 8th Graders



Principals make a difference for team progress. SAM team success in leading inquiry-based school reform depends on their principal's support of the team's inquiry work and authorization of their leadership in the school. SAM II schools and New Visions PSO schools have varied widely on these conditions for team success. In some settings, the principal has been an active supporter of the team's inquiry work and leadership; in others, the principal has pulled the team's time away from inquiry or undercut their decisions. Principals also vary in the stance they take on the program's credentialing function and the likelihood that strong teacher leaders will leave to be administrators in another school.

One SAM II principal expressed the positive stance needed to nurture the team's leadership development: "I liked the idea of bringing other people into the leadership role and running the school, taking on the responsibility of administrative tasks and so forth." In discussing the threat that strong teacher leaders would leave the school once certified, this principal said:

So my philosophy was: when I became assistant principal there were people there that helped *me* get [there]...they mentored me, and they encouraged me to move on to an assistant principal's job...so I think I should do the same.

Teachers on this school's team reported strong support from the principal and had significant decision authority in the school. For example, they took over leadership of the regular morning grade-level meetings of advisors and instructors and created a model for these meetings that featured systematic focus on each individual student. The team's ability to innovate with colleagues hinged on the principal's active endorsement of their leadership. Significantly, principal support did not entail leading the team's work. As one team member put it: "I really credit the principal and the assistant principal...that they really took this on themselves and said, 'Okay, they're in this to learn, and we're going to do everything we can to support them.'"

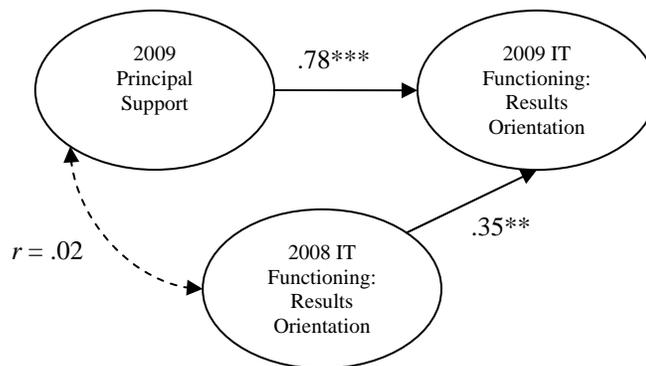
On the other end of the spectrum, team members in a SAM II Empowerment school reported that their principal actively inhibited their leadership. When asked in a focus group about principal support and involvement in their work, one team member said: “I would say that [involvement] is in the *opposite* direction [of support]. [S/he] has been *very* involved and is making decisions that [s/he] then said are *our* decisions, that weren’t necessarily our decisions.” Another team member added:

I’d say that it was kind of...a lack of organization on [principal’s] part that leads to sometimes precipitous and sometimes less than democratic decisions...[but then] moving back to the rhetoric of democratic leadership, saying ‘how are you going to move forward on your decision?’ kind of thing.

In this particular school, the principal over-rode the SAM team’s plan to work with grade level teams as a strategy to spread inquiry work and engineered a pull-out program for target students. Team members were consequently stuck with evaluating an intervention they had not designed.

Inquiry Team survey data for 2008 and 2009 capture effects of varying levels of principal support on teams’ progress on leadership for data-based improvement. Using data for New Visions schools, we find a strong positive effect of IT member ratings of principal support on their progress on leading inquiry-based reform in the school (see Figure 6). The data suggest that growth in team leadership over time is significantly influenced by a principal’s stance on distributing leadership and using SAM’s model as the engine of school improvement efforts.

Figure 6. Principal Effect on Inquiry Team Functioning (N=38 teams)



The statistical results track with interview data on ways in which some principals inhibit school teams’ progress. Absent endorsement of the team’s work and resources to support it, the inquiry team flounders.

Facilitators make a difference for team progress. SAM facilitators are fundamental to school teams’ success in the credentialing program in several ways. They bring the rich curriculum alive in seminars, give teams feedback and support on assignments, utilize SAM tools and broker additional resources to support each team’s work, scaffold and facilitate inter-visitations between schools, and help manage relationships between team and administration. In terms of the last role, it appears that a strong SAM facilitator is critical in enabling team progress

in a school where the principal might inhibit leadership. In such a setting, the facilitator can both mediate the team’s relationship with the principal and help develop her/his understanding and support of the team’s work on inquiry-based reform.

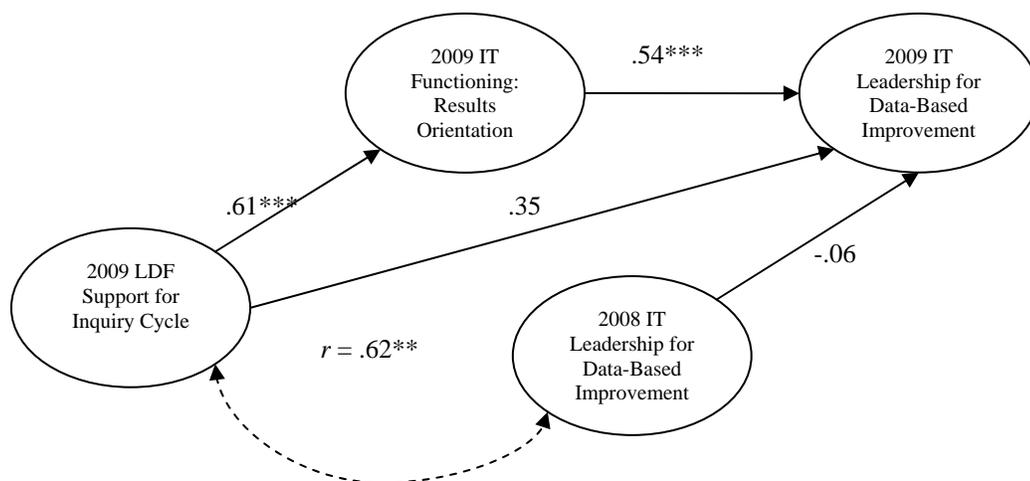
The struggling SAM II team mentioned earlier got a seasoned facilitator for their second year of the program. Participants reported a qualitative improvement in their progress. One teacher said: “She’s *amazing*... she brings such a wealth of experience and knowledge that anything she says it’s like ‘yes! yes!’” Another said:

It [role of facilitator] *definitely* has changed. These past couple months...have been really great because she works with us not only on the things that we’re doing for SAM and helping us navigate through – “okay, how are we going to be able to truly impact student achievement and get around all of the teacher mishmash that sometimes happens, or the scheduling mess that sometimes happens?” – but she also works on developing the administrative leadership pieces...of the program.

Another team member commented: “Supportive, yeah. And also focused on the research...But I also feel like she’s taken on a *much* larger role than that in our school.” In this case, the SAM instructor was facilitating change in the school’s leadership culture as well as supporting the team’s work in the program.

Data from the 2008 and 2009 IT survey in New Visions schools support the claim that facilitators are pivotal to the development of a well-functioning inquiry team that then leads inquiry-based reform in the school. Controlling for baseline levels of IT leadership, we find statistically significant effects on later measures of IT leadership of a team’s ratings of Leadership Development Facilitator (LDF) support for the team’s inquiry practice (see Figure 7).

Figure 7. Facilitator Effects on Inquiry Team Progress (N=38 teams)



These data suggests that SAM facilitators are critical agents in developing school teams' leadership of inquiry-based reform in their schools. Evidence of their variable success raises the question of how facilitators learn to strategically support SAM work in particular schools.

How SAM works up close: Case Illustration

A case study of SAM work in Marble Hill High School for International Studies (Marble Hill) illustrates how SAM teams carry out an inquiry cycle, define needs for system change, and lead culture change in the school. We follow the work of two successive SAM teams over the course of four years – incorporating target student data the teams reported in their assignments and presentations – and then examine outcomes for school leadership and culture change, as well as student outcomes beyond those of the target group.

Marble Hill was part of a cohort of five SAM II schools in former Region 1 and part of a cohort of seven SAM III schools in New Visions PSO.²¹ The school has been successful in both transitioning SAM II graduates into administrative and leadership positions and in spreading and sustaining inquiry work. All five SAM II participants earned their credentials and three of the five still work at the school. One became principal when the founding principal retired and another became an AP. SAM III participants – a math teacher and an ESL teacher – will graduate and receive their credentials in 2009-10, and a SAM IIIB cohort of teachers is beginning the program in 2009-2010.

Our case analysis takes each team's work as a stage of the school's movement to deepen its culture of assessment and broaden inquiry leadership, as well as to expand the sphere of student success.

School context. Marble Hill is a small college preparatory school located in the Bronx. It was founded in 2002 with a focus on international connections, global awareness, and language learning. Students and staff speak over 35 languages. The school's mission is to promote understanding and knowledge of other cultures, and its program includes community service, inquiry-based learning, and required four years of Math, Science and a minimum of three years of second language instruction. Portfolio assessments are part of school tradition, and all students submit two portfolios each year. A senior exit project includes a semester class with the following components: college research, college application, career planning, research paper, community service reflection, and oral presentation in the last semester of the senior year. The school graduated its fourth class in 2008-09.

Marble Hill serves about 430 students in grades 9-12. In 2007-08 the student body was 17% African American, 60% Hispanic, 6% White, and 15% Asian; 34% of the students were English language learners (ELLs) from a wide range of cultural backgrounds, and 1% were classified as special education students. The school receives Title 1 funding, and 85% of students are eligible for free and reduced price meals. The local community faces social and economic

²¹ A cohort of small schools convenes weekly for an evening seminar lead by a SAM instructor. Teams share and critique each other's inquiry work as well as conduct inter-visitations with other schools in the cohort.

challenges, and the school successfully helps many students overcome difficulties in their lives, progress well, and attain good grades. The average attendance rate in 2007-08 was 93%.²²

In 2008-09 Marble Hill had 31 teachers fully licensed and permanently assigned to the school; two-thirds had more than two years of experience teaching at the school and just over one-third had more than five years at the school. The majority (88%) had a Master's Degree or beyond, and 100% of core subject classes were taught by "highly qualified" teachers, as defined by NCLB. Marble Hill received A's on both 2006-07 and 2007-08 NYC DOE Progress Reports.

Prior to its involvement in the SAM program, Marble Hill had developed a strong student assessment culture through the use of student portfolios in addition to a range of standardized and formative assessments. The school's SAM work defined new frontiers of assessment practice: to focus on students who were least successful and most at risk of not graduating, to develop fine-grained assessment of these students' skill gaps, and to move the students through high-leverage interventions.

SAM II. The first SAM team included four teachers, two who were founding teachers in the school; one became principal after completing the program, when the founding principal retired, and the other became an AP after graduating. Their SAM program began in January 2006 and ended in Spring 2008. Marble Hill's principal was initially involved in co-teaching the Region 1 cohort's weekly seminars, along with other principals as required by SAM's design (after the first year SAM instructors were solely responsible for the seminars). The principal also supported the team's inquiry work in the school – initially taking a directive role and then stepping back – and actively promoted a vision for the school of using evidence to meet the instructional needs of all students.

Team members developed a close and productive relationship with one another and valued the SAM seminar and facilitator support of their work. One participant commented that she [facilitator] "gave us a chance...to reflect at the end of every class. And she reacted to the reflections." The team also valued presenting their inquiry work to other school teams during seminars and receiving their feedback, as well as learning from the work of other teams. Assignments and the "real" work of SAM were carried out in the school, during common planning time and beyond normal working hours.

This team began its inquiry cycle by analyzing 11th and 12th graders' transcripts in order to identify a target population. They examined Regent scores and credit accumulation to determine patterns and correlations between passing rates in subject area classes and achievement on Regents examinations. Results showed somewhat lower passing rates in English and Global History Regents for 11th grade ELLs, compared to other students. The team focused on these two examinations – each of which required a high skill level in academic written English – and selected a target population of 36 ELLs in 10th grade who appeared at risk of failing the exams based on their performance on Mock Regents (scores below 65). Transcript analysis revealed that the target students generally were under-credited in Social Studies and

²² Data from NYC DOE and NY State Progress Reports, Quality Reviews, Learning Environment Surveys, Accountability and Overview Reports, and Comprehensive Educational Plans (2007-08). Note: these student data coincide with the SAM II team's third year and thus attendance rates partly capture outcomes of their work.

English, were former Students with Interrupted Formal Education (SIFE), had recently arrived in the US, were over-aged for their grade level, and entered school mid-year.

Inquiry to determine a sub-skill and learning target involved the team in item analysis of the 10th Grade Mock Global Regents. Results led the team to focus on ‘critical thinking skills in the use of written English’. Team members appreciated the facilitator’s timely and relevant feedback as they worked to refine their focus and to develop assessments and interventions. The interventions they designed were to be implemented the following school year (2006-07, the second year of SAM II): a mandatory after-school tutoring program and plans to customize target students’ programs to meet their individual academic needs.

During the summer of 2006, the team participated in SAM’s two-week Summer Intensive session, which featured readings on Systems Thinking and team work to design leadership strategies to move inquiry across the school. In examining possible leverage points within the school, the team identified information flow and programming as crucial. They determined that, although teachers had access to information, there was no structure in place for sharing student data among teachers across grades and subject areas; nor was there systematic use of formative assessment to guide instruction. The team took on these challenges of improving the quality and use of information on student performance. Further, they decided to take advantage of the school’s ability to personalize student programs as a way of addressing the ELLs’ needs for additional English classes and seminars in specific Regents areas. The team’s agenda for leveraging and leading inquiry-based improvements addressed particular needs of its school and target students.

During their second year in the program, SAM II participants kept their focus on Global Studies and ELA Regents. However, based on feedback provided from other classroom teachers, the team realized that the original focus of “critical thinking in writing” might be too broad. Conversations with colleagues about students’ class performance led to the conclusion that they struggled most with “main idea” and “detail” questions. SAM participants then refined their focus on critical thinking to hone in on “inferencing” and “paraphrasing” as learning targets. According to team members, “we went more granular.” Their further research and success in “going small” with the data and learning targets ultimately led to an exclusive focus on “main idea and detail in multiple choice questions.” One team member explained:

We changed our skill three times, or at least what we were really focusing on. And it’s because every time we tried something we thought we needed to go even deeper than that, even smaller than that. So I think that just that *realization* was a change every time. And it was something that [our facilitator] encouraged.²³

The principal supported the team’s work by meeting with them weekly and providing input on all initiatives. Along with the facilitator, she urged the team to hone in on something specific for skills and sub-skills: “You have to find out where a student is stuck and you have to keep looking until you find it.” The facilitator helped the team make sense of the inquiry process

²³ One lesson from SAM II was that teams needed to “go small in order to go big.” In SAM III, teams were instructed and guided to identify skill gaps, then sub-skills, and finally “learning targets” that are granular enough to be teachable and move students’.

and its core principles. According to team members, she was honest about learning alongside them and helped to build trust within the team and to create a space within which she they could make sense of inquiry together.

SAM participants found conversations with their school colleagues about target students' skills and learning targets to be helpful and non-threatening to the other teachers; however, they encountered initial resistance to the LITs required by the program. Although the transcripts were being used solely for SAM participants to see target students' learning opportunities and classroom behavior, teachers were at first reluctant to open their doors. Their initial concern that they might be evaluated by team members, who were pursuing their administrative credential, lessened after the SAM team presented work at school-wide professional development sessions and made clear that their focus was on target students' learning. This sharing, coupled with the principal's legitimization and support of inquiry, facilitated team members' access to other teachers' classrooms and their ability to assess how well the curriculum and classes supported their target students' particular learning needs.

SAM team members became increasingly skilled in working with colleagues on analyzing student performance. Before conducting a LIT, a team member sat with the teacher to examine student data, and afterwards they shared and discussed the transcript. Not only did teachers become more comfortable having a team member in their classroom, but they often asked for advice on how to better reach target students. LITs gave teachers opportunity to see their instruction through the lens of the target students in their classrooms and helped them become reflective about their practices.

Over the course of two years, the team's target population dwindled in size as interventions proved successful, decreasing after the first year from 36 to 26 due to discharges and improved student performance. At this point, the team added mentoring and group meetings to the after-school tutoring, thus increasing instructional time with students. By the end of the second year of SAM, the team had surpassed its original goal of a 50% passing rate of 65 or higher for Global Regents and 43% for ELA – 76 percent and 56 percent of the students, respectively, passed the two exams. This outcome left the team with 14 remaining target students to track and mentor going into their third year in the program.

In the third and final year of SAM II (2007-08), the team aimed to continue the expanded repertoire of interventions from year 2 and to identify a new target population of students. Again the team analyzed Regents scores to identify students who had not yet achieved at 65 on Global Studies and ELA Regents. They then designed and used their own assessments to identify gaps in skills and sub-skills. During the first year of inquiry, the team had initially relied on available Acuity assessments, but “we didn't think that we got much from it...we wanted to get the students where they were...We created our own questions in which the students have to read passages and answer comprehension questions. And they were all targeting the specific skills that we wanted to find in the kids.” As in the first year, the team used their assessment every 4-6 weeks to track student progress.

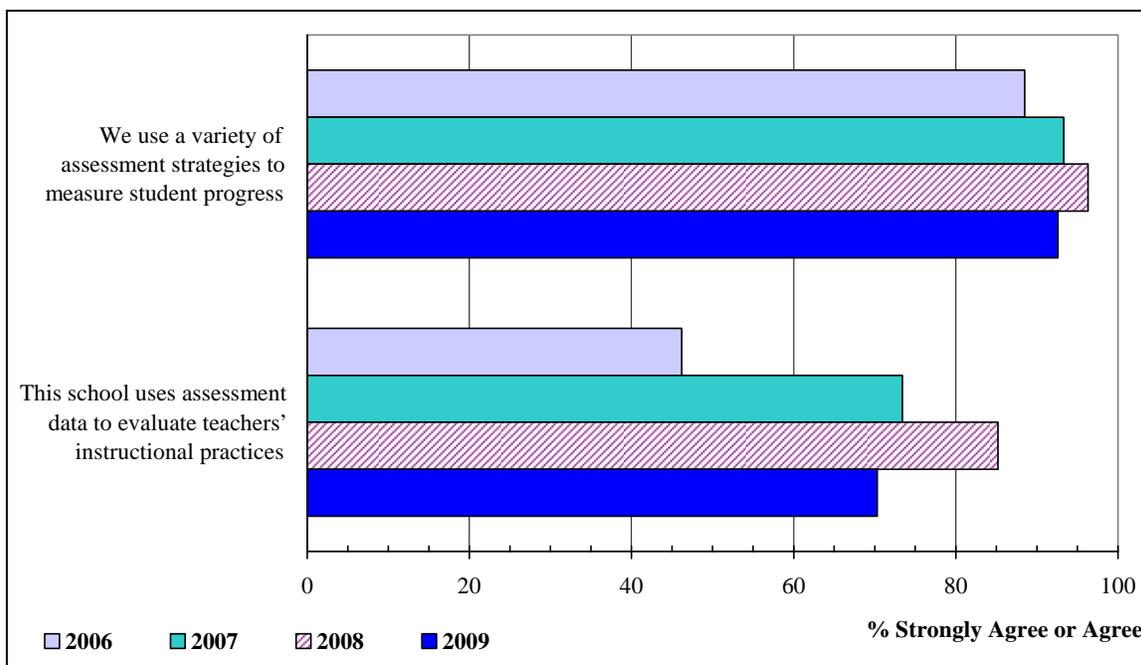
By the end of the year the team was proud that their remaining 14 target students had passed the Global and ELA Regents with 65 or higher. Also, the team's ongoing sharing of data

and professional development with colleagues had nurtured a culture of inquiry in the school. Initially SAM participants “were sharing the information with [others] in the hope that they would use the same strategies...[that were] being used in after-school by the teachers in SAM.” They found that teachers were surprised by how quickly the interventions could accelerate the students’ performance and became consistent intervention adopters. Further, students had come to take more ownership over their learning. SAM teachers noted that “students were also monitoring their own progress...That was great...making the students feel that we were there for them and we were not going to leave them alone, that this was something we were all going to do together, was great.”

The SAM II team’s work prompted the increase of common planning time and growth in teacher collaboration in the school, as well as increased attention to the link between student performance and instruction. Figure 8 shows a dramatic increase during this period of time in teachers’ reports that the school is using assessments of student performance to evaluate instruction – a change that has largely been sustained.

The SAM II team’s success in improving student outcomes and their leadership in the school developed faculty interest in and commitment to using inquiry as an improvement strategy. According to team members, “it’s not just about SAM but about the school.” Decision-making in the school became increasingly tied to evidence of student learning needs.²⁴ In effect, SAM helped to develop a school culture of decision making that uses data on student achievement and learning needs to focus and evaluate policies, programs, and practices.

Figure 8. Marble Hill Teacher Reports on Assessment Use: 2006-09



²⁴ Our 2006-2009 surveys document an increase from 52 percent to 86 percent of teachers who agreed with the statement “The faculty has an effective process for making group decisions and solving problems.”

SAM III. The second Marble Hill team began its program in February 2008, while the SAM II team was completing its third year. The new team included an ELA/ESL teacher who was in her first year at the school, having previously worked with the Marble Hill founders at their former school, and a math teacher who had been at the school for four years. Each had additional responsibilities in the school: the math teacher coordinated the Advisory Program, and the ELA teacher became the data specialist for the school's Inquiry Team and subsequently the school's ESL coordinator.

This team's program context differed from that of the school's SAM II team. Their seminar instructor was new to the SAM program, and the school had a new LDF who was Marble Hill's former principal.²⁵ The school's current principal was a SAM graduate and thus had an insider's perspective on the credentialing program and commitment to leading inquiry in the school. Further, as she noted: "little program change and little teacher turnover" contributed to a smooth start for the school year and helped to ensure continuity of ongoing inquiry work.

Building upon tested and successful SAM II methods, the new team identified a target population of 24 ELL students in the 11th grade who had failed at least two Regents examinations. The students represented 12 different countries and various language backgrounds. The team continued working with these students through their 12th grade year and added four additional students whose ELA Regents scores were below 65.

During the first year of their SAM work, the team focused on building target students' academic vocabulary. In the second year, their ongoing analysis of data from Regents examinations and Scantron assessments led them to focus on ELA, reading as a skill, and reading comprehension as the sub-skill. The team used *The Seven Habits of A Proficient Reader* as a resource for refining learning targets: main idea, inferring, summarizing, and questioning.

The SAM III team designed and implemented both programmatic and instructional interventions. Programmatic interventions included: individualized programming; a two-hour ELA block; Regents prep courses for Global Studies, Math, US History, and Living Environment; mentoring sessions every 4-6 weeks; mandated after school tutoring; Saturday Academy; and small group instruction during the school day for students most at risk in ELA and Math. Instructional interventions consisted of: intensive independent reading in ELA block; direct reading instruction in *The Seven Habits of a Proficient Reader*; strategies practice in whole class, group, and individual reading; individual vocabulary work in Global Studies and Math; instruction in learning targets using both students' independent reading level and Regents level texts; and frequent conversations between ELA, Global, US History, and Math instructors.

SAM III teachers tracked and presented Scantron data measuring their students' progress over time (see Table 2). Students showed steady improvement on these formative assessments of

²⁵ The school had joined New Visions for Public Schools' PSO in 2007-08, and New Visions assigned an LDF to each school primarily to support the work of its Inquiry Team. The former Marble Hill principal joined New Visions as an LDF after retiring; she became responsible for several small schools, including four of the SAM II schools in former Region 1. As Marble Hill's LDF, she worked with the principal and the school Inquiry Team to support its improvement efforts.

reading skills. On average over 13 months there was a 1.4 grade level improvement. One student moved six grade levels.

This team dealt with the challenges of “going small” with data, as had SAM colleagues in their and other schools. What helped most in fine-tuning their assessments and interventions were: 1) focus and reflection on links between teaching and target students’ learning in their own classroom and 2) consultation with colleagues who teach the target students to diagnose the students’ skill gaps and assess their progress.

Table 2. Scantron Results by Grade Level Equivalency in Reading, Nov. 2007-June 2009

Name	Nov-07	Feb-08	Jun-08	Oct-08	Dec-08	Apr-09	Jun-09	Gains in academic year 08-09	Overall gains
	4.8	5.9	6.1	5.1	7.4	7	8.4	3.3	3.6
	2.1	2.3	3.3	3.2	3.7	2.7	2.8	-0.4	0.7
		5.5	6.3	5.9	5.9	9.9		4	4.5
	2.5	2.5		2.5	5.1		5.5	3	3
	3.9	3.2	4.7	3.6	3.8	2.4	5.4	1.8	1.5
	5.8	4.9	6.1	6.3	7.9	9.9	6.2	-0.1	0.4
	2	<2		<2		2.9	5.2	3.2	3.2
	4	5.7	5.2	3.8	6.2	7.1	6.3	2.5	2.3
	6.6			7.2	5.2	5.9		-1.3	-0.7
	2.9	2.2	2.5	2.3	3.7	3.8	3.3	1	0.4
	3.7	3.8	4.8	5.4	4.2	7.5	5.1	-0.3	1.4
	4.5	3.5	3	3.5	2.4	4.7	3.1	-0.4	-1.4
	4.1			3.9	4.4	7	7	3.1	2.9
	6.3	4.1	7	5.3	7.4	6.8	7.4	2.1	1.1
	2.8	4.9	8.1	>9.9	9.4	6.1	9.6	-0.3	6.8
	5.8		3.7	3.3	3.3	7.2	9.9	6.6	4.1
	5	5.5	5.7	7.3	6.4	8	6.3	-1	1.3
	3.1	3.3	3.5	3.4	4.9	4.9	5.9	2.5	2.8
	4.1	5.2	4.4	4.1	6.3	5.3	5	0.9	0.9
	3.4	2.6	4.1	4.3	4.9	3.7	5.8	1.5	2.4
	2.5	2.5	3.5	4.5	2.6	2.5	5	0.5	2.5
	3.4	3.3	4.8	5.3	5.8	5.4	4.1	-1.2	0.7
Average change								1.41	2.02

In the last analysis, the SAM team was looking for student success on high-stakes Regents exams. Although not all target students passed the Regents, “we had quite a few kids who on their English Regents went from getting something like 25 points to getting 55 or 53 points. So they’re really in the ballpark of passing soon. But they came up maybe 30 points! And a 30 point increase...is astounding.” The team summarized the overall pattern of target students’ movement on each Regents exam with data shown in Table 3. By June all of the students had passed the ELA exam.

Considering these results, the school’s LDF commented that the SAM III team “is producing really good work. I mean they’re really dedicated to this population of kids. And they’ve got a *lot* of kids. It’s a pretty big target population. And the stakes for them are very high because they selected kids who they were hoping to graduate.”

Table 3. Regents Performance: Number of Students’ Highest Scores in Each Category Before January 2009 Compared to After the January 2009 Regents Examinations (total=22)

	US History		Global		ELA		Math		Living Env.	
	pre Jan	post Jan	pre Jan	post Jan	pre Jan	post Jan	pre Jan	post Jan	pre Jan	post Jan
Below 55	13	10	7	4	15	7	2	0	2	1
55-64	6	7	7	8	7	10	9	4	10	7
65+	3	5	8	12	0	5	11	18	10	14

A member of the school’s Inquiry Team commented on SAM’s role in bringing staff focus and commitment to students who might otherwise fall outside the sphere of success:

If it wasn’t for SAM, a lot of these [target] kids would have fallen through the cracks. Some would have just been mis-scheduled. Students have so many little needs and that type of focused work to address these needs would not have happened without the IT process. Furthermore, now the teachers know these students [target group] really well: other people start thinking about these kids differently.

All but three of the SAM III team’s twenty-four target students graduated on time with their class.

School culture change. Over the past four years, Marble Hill has developed a school culture of inquiry. Each of the SAM teams has involved their colleagues as participants in inquiry and implementers of instructional interventions. The principal observed, “...we’ve pulled in others who aren’t in SAM, and I think that’s really helped... [The new SAM team has] helped spread that process around.” NV support staff noted of Marble Hill that “in terms of the percentage of teachers there, they’re getting into a kind of school-wide thing that’s becoming just the way that they work in their school, and they’re really into it.”

The school has institutionalized structures and norms to sustain an inquiry culture. Common planning time is established, and most teachers welcome the chance to work with colleagues to improve instruction for their students. SAM graduates and participants have been taking on leadership roles in the school, and a third cohort of three SAM participants began the program in Fall 2009. It appears that Marble Hill has moved beyond a ‘tipping point’ at which inquiry has become part of ‘the way we do things in the school’ and at which there is an irreversible collective commitment to making sure that all students succeed. The school Inquiry Team’s leadership will continue to be important for leveraging and modeling the process. This year it is focusing on struggling 10th graders and will carry out a cycle of analyzing skill gaps,

identifying learning targets, designing and evaluating instructional responses, and sharing effective strategies with all teachers in the school.

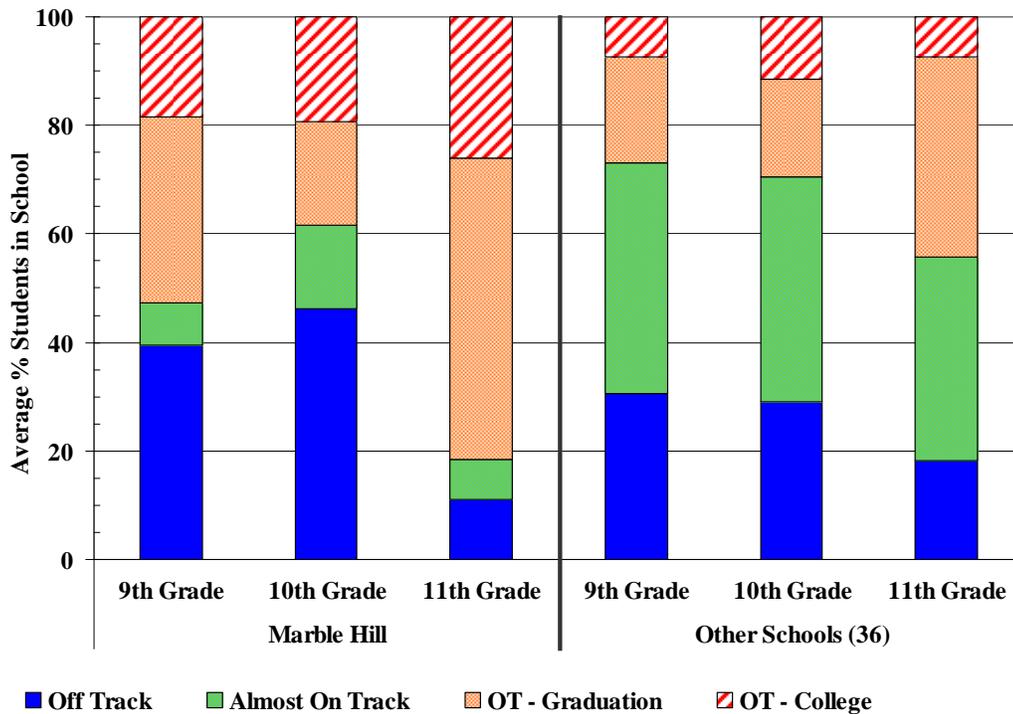
Marble Hill's progress on inquiry-based reform over the past four years is captured by trends on quantitative measures of the school's inquiry culture. For one, the school's ratings on the DOE's Quality Review, which emphasizes a school's use of data to assess and improve instruction, improved significantly. In 2006-07, Marble Hill was rated "well developed" in all areas including data inquiry and praised as being an "effective school where high quality leadership has a significant effect on the culture and learning environment for students, staff, parents and wider community." In 2007-08, Marble Hill received ratings of "outstanding" across all criteria, reflecting the spread and deepening of its inquiry work. The Quality Review stated "the school is using a wealth of data on students to provide exceptional outcomes in students' progress and performance," including "excellent assessment systems" which teachers use from the very start of the school year to determine instruction and "interim tests" to assess student learning (with a special focus on English language learners). As a result, "teachers know their students exceedingly well both academically and socially," and students (particularly ELLs) "make exceptional progress in their achievement."

Further, on our teacher survey measures of a school's culture of assessment use, Marble Hill moved significantly during its third year of SAM (Figure 8 above). The jump in 2007 can be interpreted as a qualitative shift in the way assessments were used being in the school as a function of the SAM team's leadership of inquiry with the staff. Four-year survey trends show that the school has sustained its practice of using student assessment data to evaluate instruction.

Student outcomes. Outcomes for students beyond the SAM teams' target group are reflected in the increased proportion of students on track to graduate across grade cohorts. Figure 9 shows on track measures for students in each cohort who entered the school with 8th grade ELA scores below *Proficient*. Among such students in the 11th grade in 2008-09, over 80 percent were on track to graduate or attend college – more than double the percent on track in 10th grade and about 30 percent higher than in 9th grade. Further, Marble Hill's 11th graders who entered the school with weak ELA skills far exceed the performance of similar 11th graders in the non-SAM comparison schools (over 80 percent versus 45 percent). The sharp jump between 10th and 11th grade students' performance likely reflects both SAM teams' focus on the older cohorts for their interventions. As the school's Inquiry Team shifts its focus to 10th graders this year, the pattern of improved student outcomes across grade cohorts should be more incremental.

The school's high four-year graduation rate – estimated at 95 percent for 2007-08 – is further evidence of the SAM teams' progress in bringing students into the school's sphere of success. Through strategic action to address learning needs of students who enter the school not well-prepared to learn from its enriched educational program, the school is making a difference for students who otherwise might fall through the cracks.

Figure 9. Marble Hill Student Outcomes: “On Track” Performance for Students whose 8th Grade ELA Scores were Below *Proficient*, by Student Cohort



Lessons and Issues for Further Research

Lessons from SAM to date contribute to a growing knowledge base on inquiry-based reform and frame issues for ongoing evaluation research. Evidence that schools’ sustained participation in the credentialing program develops an inquiry culture and expanded sphere of student success supports SAM’s theory of action. It also begs the question of whether and how schools can achieve these outcomes without the SAM curriculum and relatively intense facilitator support. After less than two years of inquiry work, New Visions schools not involved in “SAM-cert” have widely varying intermediate outcomes. Our evaluation will continue to track these schools’ progress and capture lessons from their struggles and successes. We also will continue to document the workings of the certification program and what facilitators and principals are learning about how to support a team’s development of inquiry and school leadership skills. Here we sketch specific lessons and issues to guide ongoing research.

SAM’s credentialing program and school reform model are synergistic. SAM takes a long-term perspective on administrator placement. It places priority on developing participants’ leadership skills so that when they do take an administrative position, most likely becoming an AP in their or another school, they will be able to make a difference for student achievement. SAM graduates have had first-hand experience leading educational improvement in urban high schools and developed skills in:

- Collaborating with colleagues as part of a school leadership team
- Using data and inquiry to focus and evaluate efforts to improve student learning
- Leading change in school professional culture toward a focus on student results and inquiry to improve instruction.

Practice-based learning through SAM assignments, supported by the program's tools and facilitators, translates into both improved leadership in the school and graduates' capacity to lead improvement as administrators in other schools.

As New Visions and New York City work to scale SAM's model for inquiry-based leadership and school reform, the question of how the credentialing program operates to promote school success is important to consider. And specifically: what facets of the SAM program can be replicated effectively without seminars, assignments, and certification?

Also important is the question of time needed for inquiry practice and leadership to take hold in a school. Schools where we documented broad, deep, and sustainable culture change had participated in the program for at least three years. And prior research indicates that school reform initiatives produce significant student outcomes only after their third year of implementation. Might we begin to see significant growth in inquiry leadership within non SAM-cert schools in New Visions and other SSOs over the next year? Evaluation results for 2009-10 and comparisons of outcome trends for new SAM-cert and non-cert schools will provide a fairer assessment of the program's effect on change.

When implemented well, SAM yields expected professional and student outcomes. SAM teams that fully implemented the model – developed a well-functioning team, carried out the cycle of inquiry with rigor, and led inquiry with colleagues – successfully moved their school culture and improved student success. Quantitative trends and case studies support SAM's theory of change. In both small high schools and in the SLCs of large restructured high schools, teams that mastered the inquiry cycle led their colleagues toward a culture of assessment use. In turn, growing proportions of students transitioned to being on track to graduate.

The trajectory of change we observed corresponds with SAM's theory of action. As a school team successfully moves target students on a particular skill, members develop a new perspective on the problem of student failure and on teachers' ability to respond in ways that promote student success. The team begins to engage colleagues in using data to identify and address student skill gaps and to lead change in school systems and culture. SAM teams' leadership took many forms – involving colleagues in instructional responses to target students' needs, sharing data about successful (and less successful) responses, offering transcripts of target students' behavior in classrooms to help shift teachers' focus toward learning, mentoring colleagues in use of data for inquiry into students' learning needs, pushing colleagues on their thinking about why students struggle, involving students in reviewing their data and developing their agency, and working with school administrators to organize and support new cohorts of inquiry teams in the school.

Although we can document the broad arc of change and ways in which teams engage their colleagues and students, we know much less about the developmental stages of change.

Within a SAM team, what qualitative shifts in individual perspective and team practice occur that enable team members to implement the model and become agents of school change? At the school level, through what stages does the professional culture move to reach a tipping point where inquiry norms overcome teacher resistance to change? Answers to these questions would provide a knowledge base to help focus strategic, effective approaches to facilitating inquiry-based reform.

Implementing SAM requires high-quality facilitator support. SAM research as well as broader literature on instructional improvement initiatives provide evidence that external coaches or facilitators can be key agents in leveraging and supporting change in professional practice. This is because routines and habits of mind in teaching, as well as typical instructional and teacher assignment policies in schools, inhibit change. Implementing SAM's inquiry model entails a significant challenge to teachers' and administrators' thinking about whether and how they can improve the success of struggling students. Many see students as victims of poverty and dysfunctional families and feel helpless to make a difference. A skilled and trusted coach can create the disequilibrium essential for individuals to change their mind about why students fail in the school and how they can meet their learning needs. Once teachers see that students can learn and improve their performance, they become invested in the work of SAM.

Significant quantitative effects of facilitator support on Inquiry Teams' progress indicate that *variation* in the quality of facilitation across New Visions schools makes a difference for school teams' progress. Unless a facilitator has deep understanding of principles of inquiry and ways in which it challenges team members' thinking, s/he will not be prepared to leverage and support shifts in team members' thinking. Ensuring that structures and routines are in place is not sufficient to move a team beyond ritual practice to serious inquiry work and leadership in the school.

We need to know much more about the ways in which a facilitator brings about qualitative shifts in a team's inquiry practice and enables members to overcome typical roadblocks to the next developmental stage. In addition to research on developmental stages of inquiry team and school culture change, Research focused on strategic facilitator moves that support change at different stages would complement that on developmental stages of inquiry team practice and school culture change. Results would contribute important knowledge for facilitator practice – to help guide their decisions, for example, about when to create disequilibrium in a team's thinking and when to help teams consolidate new thinking and practice. Further, we need to better understand how facilitators learn to use and refine such. How much and what kinds of investment in facilitator development are needed to scale up SAM? Over the next year, the evaluation will be documenting SAM facilitators' learning about both development stages of their teams' work and effective facilitation strategies to support their teams' development.

Principals make a difference for an Inquiry Team's progress. SAM's theory of action calls upon the principal to build a team of staff leaders to pioneer and ultimately lead evidence-based practice in the school. In SAM, a principal's leadership centers on creating ample time for teamwork, supporting the team's access to and use of data on individual student performance, endorsing teachers' inquiry work as important and central to the school's improvement efforts,

and authorizing the team’s leadership with colleagues. By enabling and promoting Inquiry Teams’ work, principals help broaden school leadership and shift the focus of instructional improvement efforts from adults to students, from teacher evaluation and professional development to student assessment and instruction targeting their skill development.

Our finding that Inquiry Teams’ ratings of their principal’s support vary widely across New Visions schools suggest that some principals either do not understand the leadership roles needed to support SAM or do not buy into the model. Nonetheless, a principal’s support rating is a significant predictor of the inquiry team’s progress.

Part of this variation may come from competing paradigms of instructional leadership. SAM challenges principals to move away from a paradigm that defines principal leadership in terms of holding teachers to high standards of content instruction and regularly monitoring their classroom practice, e.g., to evaluate fidelity to an instructional pacing guide. SAM calls for a new paradigm of principal “learning leadership”²⁶ – one in which school leaders hold teachers accountable to identify student learning needs and develop high-leverage strategies to address them.

Research on how principals learn to develop new “learning leadership” practices in support of inquiry-based reform would help principals, SSOs, and SAM program leaders to implement and spread SAM. What kinds of evidence or experience make a difference in principals’ commitment to this improvement strategy over others? What kinds of principal supports are most critical at each stage of development of team inquiry practice? How does a principal learn to make strategic decisions to advance both inquiry and school staff’s capacity to address the pressing learning needs of struggling students?

System investments set capacity for scaling up SAM. An additional lesson that might be drawn from evidence that both facilitators and principals are critical to teams’ progress is that school systems play an important role in developing professional capacity to scale SAM in NYC and elsewhere. The learning curve for principals and for individuals who become SAM facilitators is quite steep because often it entails *unlearning* prior approaches to leading instructional improvement. Just as SAM teams needed to move beyond their comfort zone to develop new perspectives and practices for addressing student learning needs, leaders of change in this direction also need to reframe their roles and develop new skills to promote inquiry-based reform.

Pressing questions in the NYC context center on how much and what kinds of resources are needed to develop facilitators’ and principals’ understanding and skills to lead school change toward inquiry practices to improve student learning? Evidence from the SAM certification program suggests that investment in facilitator training needs to be substantial. The program uses weekly all-day facilitator training sessions for include curriculum development, calibrating standards for evaluating SAM team products, and collective problem solving around particular instructional challenges. The weekly sessions ensure quality control in the program and a high-functioning learning community to support effective facilitation. It establishes conditions for an effective leaning environment –one that is focused on content, learners, assessment and

²⁶ Dufours, R. and R. J. Marzano, “High-leverage strategies for principal leadership.” *Educational Leadership*, February 2009: 62-68.

feedback, and community.²⁷ Is this level of investment in facilitator development essential and scalable? What if any “short cuts” would enable the same level of success?

Future research to guide system investments would consider valuable learning resources, as well as knowledge gaps, experienced by facilitators in non-cert schools. For example, to what extent and under what conditions can facilitators with less training learn through using the SAM curricula and guidelines developed in the certification program? How can an SSO or school network utilize the SAM tools, as well as lessons from SAM practice and evaluation research, to scaffold the development of effective facilitation across schools? What kinds of ongoing learning opportunities are essential to success?

New Visions has been a leader in developing new schools around design principles, pioneering SAM, and using school inter-visitations as a way to expand schools’ horizons and cross-fertilize knowledge from practice. The PSO is well-positioned to develop design principles and guidelines for developing school administrators’ and facilitators’ capacity to lead inquiry-based school reform. In turn, this work will contribute knowledge to the broader system about how to build professional capacity for inquiry-based school reform.

²⁷ For elaboration of conditions of effective learning environments, distilled from cumulative research findings across several disciplines, see Bransford, J., A. Brown, and Cocking, *How People Learn*. National Research Council, 1999.

References

- Archer, J. (2005, December). Some Ed.D. programs adopting practical approach. Team efforts supplant individual research by aspiring school leaders. *Education Week*, 25(15), 8-10.
- Bell, R., Garcia-Vanderhorst, W., Larson, K., & Lowes, S. (2008). *SAM II Team action research report: Years 1-3*. New York: Marble Hill School for International Studies and New Visions for Public Schools.
- Bransford, J., Brown, A., & Cocking R. (2000). *How people learn: Brain, mind, experience, and school*. National Research Council.
- 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.
- Dingman, W., & Kinney, C. (2009). SAM Module 3: Cycle 6 Product. New York: Marble Hill School for International Studies.
- 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.
- DuFour, R., & Marzano, R.J. (2009). High-leverage strategies for principal leadership. *Educational Leadership*, 66(5), 62-68.
- Fullan, M. (2005). *Leadership and sustainability: System thinkers in action*. Thousand Oaks: Corwin Press.
- Fullan, M. (2001). *Leading in a culture of change: Being effective in complex times*. San Francisco: Jossey-Bass.
- Gewirtzman, L. (2009). Educational leadership program enhancement project proposal (to New York State Education Department). School of Public Affairs, Baruch College, CUNY.
- Gladwell, M. (2008, December 15). Most likely to succeed: How do we hire when we can't tell who's right for the job? *New Yorker*.
- Harris, A. (2008). *Distributed leadership in schools: Developing the leaders of tomorrow*. London: Routledge Falmer.
- 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.

- Kane, T.J., Rockoff, J.E., & Staiger, D.O. (2006). What does certification tell us about teacher effectiveness? Evidence from New York City. Retrieved on September 4, 2009 from <http://www.dartmouth.edu/~dstaiger/Papers/nyc%20fellows%20march%202006.pdf>.
- Lachat, M., & Smith, S. (2005). Practices that support data use in urban high schools. *Journal of Education for Students Placed at Risk*, 10(3), 333-349.
- Levine, M. (2005). Putting the world into our classrooms. New York: Progressive Policy Institute.
- Little, J.W. (1982). Norms of collegiality and experimentation: Workplace conditions of school success. *American Education Research Journal*, 19, 325-340.
- McLaughlin, M., & Mitra, D. (2003). The cycle of inquiry as the engine of school reform: Lessons from the Bay Area School Reform Collaborative. Stanford University: Center for Research on the Context of Teaching.
- McLaughlin, M., & Talbert, J.E. (2001). *Professional communities and the work of high school teaching*. Chicago: The University of Chicago 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.
- New York City Department of Education (2009). *Marble Hill High School for International Studies: School Comprehensive Educational Plan (CEP)*. New York: NYC DOE.
- New York City Department of Education (2008). *Progress Report: Marble Hill High School for International Studies*. New York: NYC DOE.
- New York City Department of Education (2008). *Quality Review Report: Marble Hill High School for International Studies*. New York: NYC DOE.
- New York State Education Department (2008). *The New York State School Report Card: Accountability and Overview Report: Marble Hill High School for International Studies*. Albany: Office of Information and Reporting Services.
- New York State Education Department (2008). *The New York State School Report Card: Comprehensive Information Report: Marble Hill High School for International Studies*. Albany: Office of Information and Reporting Services.

- New York City Department of Education (2007). *Marble Hill High School for International Studies: Learning Environment Survey*. New York: NYC DOE.
- New York City Department of Education (2007). *Progress Report: Marble Hill High School for International Studies*. New York: NYC DOE.
- New York City Department of Education (2007). *Quality Review Report: Marble Hill High School for International Studies*. NYC DOE.
- No Child Left Behind*. (2002). Public Law 107-110.
- Popham, W. J. (2001). *The Truth About Testing: An Educator's Call to Action*. American Society of Curriculum and Development.
- Rockoff, J.E., Jacob, B., Kane, T.J., & Staiger, D. (2008). Can you recognize an effective teacher when you recruit one? National Bureau of Economic Research Working Paper No. w14485.
- Senge, P. (1999). *The dance of change*. London: Random House, Inc.
- Sharratt, L., & Fullan, M. (2005). *The school district that did the right things right*. Annenberg Institute for School Reform at Brown University.
- Siskin, L.S. (1994). *Realms of knowledge: Academic departments in secondary schools*. Abingdon: Routledge Falmer.
- Spillane, J. (2006). *Distributed leadership*. San Francisco: Jossey-Bass.
- Supovitz, J., & Klein, V. (2003). *Mapping a course for improved student learning: How innovative schools systematically use student performance data to guide improvement*. Philadelphia: Consortium for Policy Research in Education.
- Talbert, J.E. (2002). Professionalism and politics in high school teaching reform. *Journal of Educational Change*, 3(3-4), 339-363.
- Talbert, J.E., & Scharff, N. (2008). Leading school improvement with data: A theory of action to extend the sphere of student success. Paper presented at the American Educational Research Association Annual Meeting, New York City, March 2008.
- Talbert, J.E., & Wood, A. (2007). Evaluation of BASRC Phase II: Evidence-based system reform: Outcomes, challenges, promising practices. Stanford University: Center for Research on the Context of Teaching.

Appendix A

NV-PSO Inquiry Team Survey: Scale Definitions

These survey scales were developed with data from the SAM Evaluation’s Inquiry Team Survey. The survey was administered in Spring 2008 and Spring 2009 to IT members in all schools that are part of New Visions for Public Schools’ PSO (NV-PSO) in New York City (N= 187 from 57 schools in 2008; (N = 291 from 71 schools in 2009). Principal components analysis was used to identify survey items that loaded on a common factor. The alpha coefficient indicates internal consistency of each scale and is shown for both 2008 and 2009. This document includes just those scales that are included in CRC’s October 2009 report.

I. SCHOOL INQUIRY CULTURE AND LEADERSHIP

➤ ***Culture of Assessment Use (2 items. Alphas = .82 & .81)***

5-point Likert scale, ranging from 1 (“Strongly Disagree”) to 5 (“Strongly Agree”)

*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)?*

	2008	2009
We use a variety of assessment strategies to measure student progress	3c	3c
We use assessment data to evaluate our curriculum and instructional practices	3d	3d

➤ ***Leadership in School or SLC: Data-based Improvement (3 items. Alphas = .89 & .93)***
5-point Likert-type frequency scale, ranging from 1 (“Never”) to 5 (“Always”)

*Now consider **leadership in your school or SLC**. Please indicate the extent to which leader(s) do each of the following...*

School / SLC leaders...

	2008	2009
Use data to identify patterns to inform decision making	4e	4d
Use objective evidence to identify, frame and solve problems	4g	4e
Use data to evaluate the effectiveness of decisions	4h	4f

II. INQUIRY TEAM WORK IN THE SCHOOL

A. Practices Linked to SAM Inquiry Team Standards

- ***Inquiry Team Performance Standards: Results Orientation (3 items. Alphas = .90 & .90)***
5-point Likert-type frequency scale, ranging from 1 (“Never”) to 5 (“Always”)

*This question concerns how you and others in your Inquiry Team **work together**. Please indicate the extent to which the team operates in each of the following ways.*

Our Inquiry Team members ...

	2008	2009
Establish clear and unambiguous measurements for assessing our success	11n	8n
Stay focused on results in the face of distractions and competing priorities	11o	8o
Willingly make sacrifices for the good of the team and the achievement of our goals	11p	8p

- ***IT Leadership of Data-based Improvement (3 / 2 items. Alphas = .90 & .90)***
5-point Likert-type frequency scale, ranging from 1 (“Strongly Disagree”) to 5 (“Strongly Agree”)

*Now consider **how the Inquiry Team works with others in your school**. Please indicate how well each of these statements describes your work.*

On the whole, our Inquiry Team ...

	2008	2009
Uses data to identify patterns to inform decision making	12b	12b
Uses objective evidence to identify, frame and solve problems	12d	12d
Uses data to evaluate the effectiveness of decisions	12e	--

B. Facilitator and Principal Support of IT Work

- ***LDF Support of Inquiry Cycle (3 items. Alphas = .94 & .92)***
5-point Likert scale, ranging from 1 (“Not at all Valuable”) to 5 (“Extremely Valuable”)

Please indicate whether or not your LDF has worked with your Inquiry Team this year on each of the following activities. If yes, please rate how valuable it has been for your leadership development.

	2008	2009
Use data to identify target students	13a	9a
Use data to identify skill gaps for target students	13b	9b
Conduct low-inference observations of classroom(s)	13c	9c
Use data to evaluate our curriculum and instruction	13d	9d
Decide on intervention(s) for target students	13e	9e

> **LDF Support of IT Leadership (4 items. Alphas = .93 & .86)**

Engage faculty in problem-solving instructional issues	13g	9g
Lead a school-wide focus on learning	13h	9h

➤ **Facilitator Standards-based Practice (4 items. Alphas = -- & .90)**
5-point Likert scale, ranging from 1 (“Strongly Disagree”) to 5 (“Strongly Agree”)

*To what extent does each of the following statements capture your **Inquiry Team’s** experience with facilitator and principal support 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.]

	2008	2009
Facilitator conveys clear objectives and expectations for our work	--	11a
Facilitator creates structures for feedback and self-assessment on our behavior	--	11c
Facilitator pushes us to think in new ways	--	11e
Facilitator holds us to the performance standards for inquiry teams, specifically, moving the students	--	11o

➤ **Principal Support of Inquiry Team (3 items. Alphas = -- & .88)**
5-point Likert scale, ranging from 1 (“Strongly Disagree”) to 5 (“Strongly Agree”)

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

	2008	2009
Principal establishes conditions for trust and open communication	--	11b
Principal actively supports our risk-taking	--	11f
Principal uses authority to push our learning in the service of target students and targeted learning goals	--	11h