Getting to Results: Student Outcomes in New and Redesigned High Schools

Evaluation of the Bill and Melinda Gates Foundation's High School Grants

Prepared by:

The American Institutes for Research

SRI International

American Institutes for Research®

SRI International®

Getting to Results: Student Outcomes in New and Redesigned High Schools

July 2005

Prepared for:

The Bill & Melinda Gates Foundation
P.O. Box 23350
Seattle, WA 98102

Prepared by:

The National Evaluation of High School Transformation
http://www.air.org/publications/pubs_ehd_school_reform.aspx

American Institutes for Research
1000 Thomas Jefferson St., NW
Washington, DC 20007
http://www.air.org

SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025
http://ctl.sri.com

The *National Evaluation of High School Transformation* is a collaborative effort between the American Institutes for Research and SRI International. This work, which began in 2001, is supported through funding by the Bill & Melinda Gates Foundation.

Getting to Results: Early Student Outcomes in New and Redesigned High Schools is part of an ongoing series of reports based on the evaluation of the Bill & Melinda Gates Foundation's high school grants. The views, findings, conclusions, and recommendations expressed herein are those of the authors and do not necessarily express the viewpoint of the foundation.

Authors:

David Rhodes, Becky Smerdon, Winona Burt, Aimee Evan, Ben Martinez, and Barbara Means

Principal Investigators:

Becky Smerdon
American Institutes for Research

Barbara Means
SRI International

Achievement Study Team:

Andrea Berger
Winona Burt
Aimee Evan
Ashley Fedor
Bridget Goosby
Matt Hickey
Stacey Merola
Amber Noel
David Rhodes
Scott Richman
Alison Schmauch
Matt Segal
Mengli Song
Tyler Vogel

Contents

Executive Summary 1

Introduction 7
- Background 8
- Research Questions 13
- Organization of the Report 14

Study Design 15
- Study Methods 15
  - Research Question 1 15
  - Research Question 2 18

Results 21
- Research Question 1: How do foundation-supported schools compare with other schools in the same district in terms of the students they serve and certain student outcomes (e.g., attendance, achievement)? 21
  - Student Enrollment 21
  - Student Outcomes: Achievement 23
  - Student Outcomes: Behaviors 26
  - Summary 28
- Research Question 2: What are some of the factors (e.g., organizational characteristics of schools and teaching practices) associated with relative success on student outcomes? 28
  - Student Attitudes 29
  - Student Achievement 33
  - Summary 38

Implications for the Initiative 39

References 43

Endnotes 45

Appendix A 47

Appendix B 51

Appendix C 65

Appendix D 71

Figures

Figure 1 Relationship Between Student Attitudes and Implementation 30

Figure 2 Effect of Reform-oriented Teaching on Student Attitude 31

Figure 3 Effect of conventional Instruction on Student Attitudes 33

Figure 4 Student Achievement and Implementation 34

Figure 5 Student Achievement and Student Engagement 37

Executive Summary

Only 34 percent of public high school students are being *minimally* prepared for college — that is, earning a regular diploma, completing a minimum set of course requirements, and reading at a basic level (Greene & Winters, 2002). Partnering with state education agencies, school districts, and other educational organizations, the Bill & Melinda Gates Foundation is working to improve American high schools by promoting the new three R's — rigor, relevance, and relationships. The foundation believes that high schools need to become places that combine rigor in the academic program of every student (not just those in an honors track) with relevance to their interests and potential career opportunities, supported by positive relationships that can inspire students both academically and personally.

A team of researchers from the American Institutes for Research (AIR) and SRI International (SRI) have been conducting a national evaluation of the foundation's initiative since 2001. This report examines student outcomes in foundation-supported schools during the early years of the initiative and draws upon a number of different data sources including: extant district demographic and achievement data, surveys of teachers and students, and site visits.

The foundation is supporting two strategies for improving student outcomes: some foundation-supported organizations are creating brand new high schools, while others are redesigning existing large high schools into effective learning communities that share a building or campus.

Combining survey and site visit data collected from schools with the student demographic and student outcome information maintained by school districts, we investigate two research questions:

1. How do foundation-supported schools compare with other schools in the same district in terms of the students they serve and key student outcomes (e.g., attendance, achievement)?
2. What are some of the factors (e.g., organizational characteristics of schools and teaching practices) associated with relative success in terms of student outcomes?

To address our first research question, we compared foundation-supported schools to the other schools located in the same district. Based on these within district comparisons we found:

- ***Foundation-supported schools are enrolling students from traditionally underserved populations.*** Looking at existing high schools in four urban school districts, we found that foundation-supported schools

enrolled a higher proportion of students who were eligible for free or reduced-priced lunch and who were members of a race/ethnic minority group than did other high schools in the same district. Foundation-supported schools generally enrolled a greater percentage of students with special education or language acquisition needs as well. Additionally, foundation-supported schools generally enrolled students who began high school academically behind students attending other schools in the same district.

- ***Patterns in student-level achievement data suggest promise for reading/ELA achievement, but not for mathematics.*** In two of the three districts where trends on state assessment data could be examined, we saw larger improvements in reading/ELA achievement over time in foundation-supported schools than elsewhere in the district. The third district experienced improvements to reading/ELA on a par with the rest of the district. The size of the gain in reading was larger in the district with new high schools than it was in the district with redesigned high schools.

- ***Trends in student-level achievement data for math were mixed.*** In on of the two districts where trends on state assessment data could be examined, we saw moderately larger improvements in math over time in foundation-supported schools than elsewhere in the district. The other district experienced moderately smaller improvements in math.

- ***Attendance rates were higher at new schools, but lower in redesigned schools, compared to other schools in the same district.*** As a group, new high schools enjoyed higher attendance rates than the other schools in both of the districts with attendance data available. Redesigned schools, on the other hand, had poorer attendance than the other schools in their districts. This finding does not mean that new schools are necessarily more successful in promoting attendance. The decision to redesign a high school is based in part on poor performance in the past. Therefore, many redesigned high schools are dealing with a legacy of problematic student behavior.

Our second question was "what are some of the factors associated with relative success in terms of student outcomes?" To answer this we examined the relationships of survey based measures of student attitudes, school attributes, and teaching practices with each other and with district-provided student outcome measures.

We found that:

- ***Implementation and student-centered instruction were positively related to student attitudes.*** In schools with higher levels of the foundation's attributes of effective schools in place, we found students more engaged, more persistent in their learning, and more satisfied with their academic progress. Students had more positive attitudes also when a greater degree of student-centered instruction was used in the classroom. The higher the level of teacher-directed instruction in the classroom, the less positive student attitudes were. These analyses controlled for student background characteristics, a school risk index, and the year of survey administration.
- ***Implementation and student engagement interest were positively related to school-level English language arts performance.*** We found a significantly higher level of implementation of the foundation's school attributes in those foundation-supported high schools that scored above their district's mean on reading/ELA tests than in those that scored below the district mean.
- ***Implementation and student engagement interest were not positively related to school-level mathematics performance.*** In fact, all the associations we observed were negative. That is, implementation levels, reform instruction, and positive student attitudes tend to be higher in foundation-supported high schools that scored *below* their district mean on math assessments than in schools that scored above their district's mean. None of these negative relationships attained statistical significance, however.

Implications

These findings have several implications:

- ***Judgments of secondary schools' performance should take into account the differential levels of academic and attitudinal preparation students bring to high school.*** Foundation-supported high schools are successfully enrolling the populations targeted by the initiative, and in doing so they are enrolling high proportions of high-need, low-achieving students. The foundation may want to focus as much as possible on "value-added" definitions of school performance. At the school-design stage, there is a need for decision tools to help design teams think carefully about the match between their organizational capacity and the needs and aspirations of the students they are recruiting.

- ***Schools need more support around mathematics content and instruction.*** In this report, we found no evidence of foundation-supported schools making progress in mathematics achievement. In part, this lack of progress may reflect some foundation-supported schools' stress on individualized learning programs built around student interests that sometimes treat mathematics in a cursory manner. Other potential contributing factors have been described in earlier reports: Some foundation-supported schools have struggled with hiring and retaining qualified math teachers, and several schools have had to build programs with teachers who were ill-prepared to teach math. Teachers have reported that good mathematics curricular materials consistent with their instructional philosophies are hard to find, and multidisciplinary resources are particularly elusive, *Creating Cultures for Learning: Supportive Relationships in New and Redesigned High Schools* (2005). Moreover, we have found wide variation in the rigor and relevance of mathematics assignments in foundation-supported schools. On average, the rigor of math assignments in foundation-supported schools was not significantly better than that of math assignments in comprehensive high schools, *Rigor, Relevance, and Results in New and Conventional High Schools* (2005). Curriculum materials, professional development, technical assistance, and coaching around mathematics content and instruction are all very much needed in these schools, as are well-qualified math teachers.

- ***It is too early in the initiative to draw definitive conclusions concerning student outcomes.*** The schools included in these analyses have been serving students for three years or less. We concentrated our analysis on only four urban districts that had a relatively large number of support-schools. Patterns may well change as the initiative and evaluation expand to additional districts, additional years' worth of data become available, and schools continue their reform efforts. School-level outcome data can be very unstable from year to year, especially in schools with small enrollments. Prior research on school improvement efforts suggests that a time frame of five to six years is appropriate for assessing the viability of an educational intervention.

Working with state education agencies, school districts, and other educational organizations, the foundation has been successful in reaching the students most in need of improved secondary education. The replication of the positive relationships reported by the guiding literature involving school attributes, classroom instruction, and student attitudes indicates that reform efforts are touching the lives of young people. Our ability to gauge the impact of reform efforts on specific student outcomes such as state assessment tests and behavioral measures is, however, limited at

this time. It is limited by the number of schools for which we have both measures of implementation and outcome data over a period of time. Many of the foundation-supported schools analyzed here are in their first and second year of existence. We collected information during the 2004–05 school year from these schools as they moved into their second and third years. We also began a new wave of schools that opened during 2004–05. As AIR®/SRI continue to collect implementation information from subsequent "waves" of supported high schools and to gather outcome data over time for existing cohorts of schools; we will be increasingly able to address these issues.

Introduction

Addressing the nation's governors in February 2005, Bill Gates declared America's high schools obsolete. Elaborating, he explained, "By obsolete, I mean that our high schools — even when they're working exactly as designed — cannot teach our kids what they need to know today"(Gates, 2005). This indictment is supported by compelling evidence. Only 34 percent of public high school students are being *minimally* prepared for college — that is, earning a regular diploma, completing a minimum set of course requirements, and reading at a basic level (Greene & Winters, 2005). Public opinion surveys reveal that most employers think high school graduates lack the basic skills today's jobs require, and the majority of workers give their high schools a grade of C or lower on how well the education they received prepared them for success on the job (American Diploma Project, 2004). Furthermore, far too many students do not even finish high school. Only 71 percent of the class of 2002 graduated with a regular diploma, and only about half of African American and Hispanic students did (Greene & Winters, 2005).

Partnering with state education agencies, school districts, and other educational organizations, the Bill & Melinda Gates Foundation is working to improve American high schools by promoting three R's — rigor, relevance, and relationships. The foundation believes that high schools need to become places that combine rigor in the academic program of every student (not just those in an honors track) with relevance to their interests and potential career opportunities, supported by positive relationships that can inspire students both academically and personally. The foundation has been making grants in support of two strategies for creating high-performing schools: some foundation-supported organizations are starting brand new schools; others are redesigning existing high schools.

A team of researchers from the American Institutes for Research (AIR®) and SRI International (SRI) have been conducting a national evaluation of the foundation's initiative since 2001. The goals and methods of the evaluation are described in detail in *Charting a Course: Evaluation Design of the National School District and Network Grants Program* (AIR®/SRI, 2003). This report examines student outcomes in foundation-supported schools during the early years of the initiative and draws upon a number of different data sources including: extant district demographic and achievement data, surveys of teachers and students, and site visits. Specifically, we begin by comparing the characteristics of students attending foundation-supported schools to characteristics of students attending other high schools in the same district to determine whether the schools supported

through this initiative are reaching the target population of low-income, historically underserved students. Next, we compare selected student outcomes (e.g., assessment scores, attendance) in the foundation-supported schools to outcomes in the other high schools in the same district. Finally, we turn to examining the relationships between the school features or attributes, teaching practices the foundation has been promoting, and students' attitudes and performance on state assessments.

This document, part of a series of reports describing our most recent findings, builds on the work of earlier reports in the series. The first evaluation report, *Creating Cultures for Learning: Supportive Relationships in New and Redesigned High Schools* (2005), examined the progress of foundation-supported schools in implementing close, supportive communities focused on learning. This report gave special attention to the development of relationships between and among students and teachers. The second report, *Rigor, Relevance, and Results in New and Conventional High Schools* (2005), investigated whether the school-level changes described in the first report have corollaries in the classroom. By defining "reform-oriented" and "conventional" teaching indices, that report examined teaching in foundation-supported schools to see whether reported school-level changes set the stage for classroom innovation and described the nature and quality of the work students produced in response to the instructional practices. Collectively, these reports provide a composite picture of what is occurring in the schools funded by this initiative. This report focuses on early outcomes for students, both in terms of test scores and other data maintained by districts and in terms of attitudes related to school and academic performance, as expressed on student surveys and focus groups.

Background

The Bill & Melinda Gates Foundation is building on its experience and recent research in adapting its strategies for high school reform. This progression has led the foundation to move toward key partnerships with state education agencies, school districts, and other organizations to improve American high schools by promoting the new three R's—rigor, relevance, and relationships. The foundation believes high schools should become places that combine rigor in the academic program of every student with relevance to their interests and potential career opportunities, supported by positive relationships that can motivate students both academically and personally. The foundation further envisions a larger policy context and K–16 system surrounding the high school, with both demands and supports designed to achieve college

The Foundation's Attributes of High-Performing Schools

Attribute	*Description*
Common Focus	Staff and students are focused on a few important goals. The school has adopted a consistent research-based instructional approach based on shared beliefs about teaching and learning. The use of time, tools, materials, and professional development activities are aligned with instruction.
High Expectations	Staff members are dedicated to helping students achieve state and local standards; students are engaged in an ambitious and rigorous course of study; and students leave school prepared for success in work, further education, and citizenship.
Personalized	The school is designed to promote sustained student relationships with adults where every student has an adult advocate and a personal plan for progress. Schools are small—no more than 600 students (fewer than 400 strongly recommended).
Respect and Responsibility	The environment is authoritative, safe, ethical, and studious. The staff teaches, models, and expects responsible behavior, and relationships are based on mutual respect.
Time to Collaborate	Staff has time to collaborate and develop skills and plans to meet the needs of all students. Parents are recognized as partners in education. Partnerships are developed with businesses to create work-based opportunities and with institutions of higher education to improve teacher preparation and induction.
Performance Based	Students are promoted to the next instructional level only when they have achieved competency. Students receive additional time and assistance when needed to achieve this competency.
Technology as a Tool	Teachers design engaging and imaginative curriculum linked to learning standards, analyze results, and have easy access to best practices and learning opportunities. Schools publish their progress to parents and engage the community in dialogue about continuous improvement.

Source: Bill & Melinda Gates Foundation. (no date). *Helping All Students Achieve* [Pamphlet]. Seattle: Author.

readiness for all students. This vision involves whole-system change, but within this system the school and its classrooms remain a location where a great deal of learning occurs.

The foundation recognizes that there are multiple models for fostering the three R's within the two basic strategies of opening new or redesigning existing high schools. By supporting a variety of high-quality schools the foundation is striving to improve graduation rates and other student outcomes for high school students, including those who have traditionally

fallen through the cracks. The foundation is supporting two strategies for improving student outcomes: some foundation-supported organizations are creating brand new high schools, while others are redesigning existing comprehensive high schools into smaller learning communities that share a building or campus. The three R's the Bill & Melinda Gates Foundation has identified are derived from a number of key attributes of effective schools that the foundation has identified through research and experience. These attributes are described in the sidebar: *common focus, high expectations, personalized, respect and responsibility, time to collaborate, performance-based,* and *technology as a tool.* Generally speaking, foundation-supported schools are expected to be inviting places, where students and adults know each other well and pursue a common mission based on high academic achievement for all students and where the professional community is collaborative and student-focused (see technical appendix for details about the construction of these measures).

The foundation has also developed a vision for the components of effective teaching and learning *active inquiry, in-depth learning,* and *performance assessment*—described in the following sidebar. The foundation calls for personalized learning environments that prompt students to take responsibility for learning, make choices, and create products that are linked to the broader community and real-world concerns. The foundation describes instructional practices that start with students' current knowledge and skills; draw on students' interests and experience; and provide students with rigorous, college preparatory learning experiences. Their instructional vision calls for collaboration between and among students and teachers and participation in community- and work-based projects. The foundation's prescriptions include clear learning goals and ongoing monitoring of progress toward those goals through multiple measures that include standardized achievement tests and performance measures such as exhibitions, portfolios, and other assessments that make visible students' understanding, reasoning, and skill levels (Vander Ark, 2001).

To develop their attributes of effective schools and classrooms, the foundation consulted a number of national experts on these topics and reviewed the research literature. Such research concludes, for example, that teacher's professional community (including *teacher collaboration*) and the presence of *common focus* of what students should be learning allows teachers to work with one another to facilitate student learning (Marks, Secada, & Doane, 1996). In their study of Chicago Public Schools, Lee and colleagues (1999) found a strong relationship between academic press (the foundation's *high expectations*) and student achievement. Additionally, evidence suggests that students who believe that they are cared for and matter (where *personalization* is high) put more effort into their schooling, which, in turn, affects their learning (Smerdon, 1999).

The Foundation's Essential Components of Teaching and Learning

Attribute	*Description*
Active Inquiry	Students are engaged in active participation, exploration, and research; activities draw out perceptions and develop understanding; students are encouraged to make decisions about their learning; and teachers use the diverse experiences of students to build effective learning experiences.
In-depth Learning	The focus is competence, not coverage. Students struggle with complex problems, explore core concepts to develop deep understanding, and apply knowledge in real-world contexts.
Performance Assessment	Clear expectations define what students should know and be able to do; students produce quality work products and present to real audiences; student work shows evidence of understanding, not just recall; assessment tasks allow students to exhibit higher-order thinking; and teachers and students set learning goals and monitor progress.

Source: Bill & Melinda Gates Foundation. (no date). *Helping All Students Achieve* [Pamphlet]. Seattle: Author.

Schools with a positive normative climate (stressing safety and orderliness, *respect and responsibility*, academic press, and students' sense of belonging) provide a strong, supportive environment that encourages students' intellectual efforts and academic achievements. Such support fosters a sense of confidence and psychological safety that allows students to ask for help, admit errors, take risks, and experience failure as they make their way along the learning path (Lee et al., 1999).

Findings from the first report in this series, *Creating Cultures for Learning: Supportive Relationships in New and Redesigned High Schools* (2005), show that in new schools exhibit the foundation's attributes at much higher degree than the large schools in our evaluation, and that new schools have many of the attributes found in more mature model schools. Although redesigned high schools are seeing slower progress as they work to change existing structures, cultures, and beliefs, they too are showing gains over their initial state, most notably in the implementation of personalized school cultures in which students feel known by their teachers and supported by them both academically and personally.

The foundation's proposals for teaching and learning are based on two decades of research on the science of learning (Bransford, Brown, & Cocking, 1999). These notions of powerful instruction are in line with a number of recent innovations in teaching — variously called authentic instruction, teaching for understanding, reform-oriented instruction, or constructivist teaching. Underlying these innovations is the notion of the

students as active learners and the teachers as guides, or coaches, in the learning process (Cohen, 1988b; Conley, 1993; McLaughlin & Talbert, 1993; NASSP, 1996; Newmann, Marks, & Gamoran, 1996; Sizer, 1992).

Findings from the *Rigor, Relevance, and Results in New and Conventional High Schools* (2005) report indicate that teaching and learning tend to lag behind structural change in foundation-supported schools — both for new schools and redesign efforts. Initially the nuts and bolts of designing and putting the small-school structure in place take precedence over curriculum and teaching. It takes time to develop the curriculum and teaching approaches being adopted by a school. Nevertheless, there is indication of more reform-oriented instruction in new schools, and some progress among redesigned schools. Key findings include:

- Teachers in new schools more frequently use reform-oriented instructional approaches than do teachers in comprehensive high schools.
- Teachers at schools that have been redesigned into smaller learning communities report doing both more reform-oriented instruction and more conventional instruction than they did two years earlier before the school's redesign.
- Teachers' language arts and mathematics assignments in new schools are more likely to have relevance, in terms of real-world connections and student choice, than the assignments given by teachers in comprehensive high schools. Language arts assignments in small new schools are more rigorous than the assignments teachers give in comprehensive high schools. Mathematics assignments in new schools also tend to be more rigorous than those in large schools, but this difference was not statistically significant.

These findings are related to an emerging body of theoretical and empirical work on student engagement; this research suggests that engagement is an essential step, according to our understanding of the process of student learning. The links among student engagement, high expectations, and improved performance have been examined by Brophy (1983), Cooper and Tom (1984), Covington (1992), Firestone and Rosenblum (1988), and Raudenbush (1984).

In reflecting on this research, we can see a number of implied causal relationships that converge to improve student outcomes — school attributes affect classroom teaching which affects student attitudes which affect student outcomes. In this report, we begin a simple, descriptive approach to examining this chain of assumptions. However, it is important to acknowledge that the data we have available are not sufficient to establish or refute the causal links in this chain. The data we have are primarily cross-sectional, and therefore we can examine only correlation

(level of association). The results presented here should be viewed as tests for the strength of the relationships the conceptual framework implies, rather than formal tests of theoretical propositions represented in the model. This is consistent with AIR®/SRI's theory of change approach to this evaluation effort. The next section identifies and discusses the research questions guiding this analysis.

Theory of Change Approach to Evaluation

Building on the work of Carol Weiss and colleagues in the 1970s, the theory of change (TOC) approach to evaluation is "a systematic and cumulative study of the links between activities, outcomes and contexts of the initiative" (Connell and Kubisch, 1998). Evaluators, employing a TOC approach, work with those planning and implementing an initiative to define an overall vision. Together they identify desired outcomes and the strategies that are intended to produce them as a basis for evaluation design.

This approach has three appealing aspects. First, it increases the likelihood that stakeholders and the evaluation team have a common understanding about the intended outcomes of the initiative, the activities that need to be implemented in order to achieve those outcomes, and the contextual factors that are likely to influence them. Second, the process of developing a TOC helps identify what outcome and implementation activities need to be measured when. Finally and perhaps most importantly, TOC evaluations examine not only outcomes but the chain of events that produce those outcomes, thereby addressing "causality" in terms of how many of the theorized effects come to pass. Although this strategy cannot eliminate all alternative explanations for a particular outcome, it aligns the key participants in the initiative with a standard of evidence that will be convincing to them.

Other evaluation approaches (e.g., experimental, quasi-experimental, and interrupted time series) tend to focus more exclusively on determining whether or not an initiative caused the desired outcome(s). These designs seek to eliminate (or account for) explanations other than the intervention that may affect the outcome(s) of interest. Once this is accomplished by randomizing the "treatment" or statistically controlling for other factors, these evaluation approaches assert causality when differences observed between "treatment" and "control" groups are unlikely to occur by chance. They do not, however, result in an understanding of how the treatment affects the outcome, or what features of the intervention are essential to its success.

Research Questions

In this report, we investigate the following two research questions:

1. How do foundation-supported schools compare with other schools in the same district in terms of the students they serve and key student outcomes (e.g., attendance, achievement)?

The primary focus of this report is student outcomes as measured during the early years of the foundation's initiative. To this end, we begin by examining the characteristics of students who are enrolled in foundation-supported schools to identify pre-existing differences in

student populations that may affect student outcomes. For example, if foundation-supported schools are skimming the highest performing students in the district, then we would expect their test scores to be high regardless of their programs. Equally important, we examine the extent to which foundation-supported schools are enrolling the populations the foundation has targeted—historically underserved students. Finally, we analyze student outcomes by comparing foundation-supported schools with other schools in their districts. To address our first research question, we draw on extant data collected on student demographics, achievement test scores, and behaviors (e.g., attendance).

2. What are some of the factors (e.g., organizational characteristics of schools and teaching practices) associated with relative success in terms of student outcomes?

This question examines the relationships, described in the previous section, among school characteristics, teaching, student attitudes, and student outcomes. Specifically, we determine whether the attributes of schools and teaching identified by the foundation are indeed related to student attitudes and outcomes and whether student attitudes are indeed related to student test scores. Are the relationships reported in the literature being replicated in foundation-supported schools? Our exploration of state assessment data focused on the subset of schools in which we could compare individual school results to overall district performance.

Organization of This Report

The remainder of this report is divided into three sections: study design, results, and implications. The following section on study design provides further details about the data and methods we use to address each of the research questions. Because these vary by research questions, we discuss these matters sequentially for each research question. After the description of methods, we separately present the results for each of the two research questions. The report concludes with a discussion of implications of these findings for the foundation's ongoing reform initiative.

Study Design

The goals and methods of the AIR/SRI evaluation are described in detail in "Charting a Course: Evaluation Design of the National School District and Network Grants Program" (AIR®/SRI, 2003). Our evaluation focuses on the goals and progress of the initiative at the national level rather than on the progress of individual schools and reform organizations. We seek to explore and test the idea that schools with the characteristics described by the foundation yield better, more equitable outcomes for students. We evaluate the degree to which reform efforts are successfully implemented through the efforts funded by the foundation and describe the factors that are key to the success of its schools.

School types and the timing of data collection efforts are shown in the sidebar. New schools were surveyed and visited during each of their first three years. The data collected allow us to track change over time in individual schools. Large schools undergoing redesign were surveyed in their planning year and again two years later. The data gathered support pre- and post-reform comparisons, as these schools were visited for three consecutive years (in their planning year, and in the first and second years of implementing the redesign). We also collected survey and site-visit data from established schools that served as models for the new small schools; these schools provide benchmark data for the new schools. To support additional comparisons with new schools, we collected data from nearby large conventional schools with comparable student populations. To examine student outcomes, we also collected extant data from foundation-supported schools and other schools within their jurisdictions (in four districts). Details of data collection and analyses are described in the technical appendix.

Study Methods

Our discussion of data sources, samples, and analytic methods is organized around the two primary research questions posed in this report.

Research Question 1: How do foundation-supported schools compare with other schools in the same district in terms of the students they serve and certain student outcomes (e.g., attendance, achievement)?

To investigate this question, we analyzed student demographic and outcome information collected from four urban districts with multiple foundation-supported high schools [1]:

- District One: Mid-Atlantic district
- District Two: Midwest district

School Types Included in This Study

This chart describes the types of schools that supplied information used in this report. Sample sizes for each type of analysis are noted throughout the report.

New schools:	New small autonomous schools that received foundation funding for their first three years.
Data:	Surveys and site visits in each of the school's first three years. Rolling sample of schools (i.e., new schools are added each year, and each stays in the sample for three years), beginning in 2001–02.
Redesigned schools:	Large comprehensive high schools receiving foundation funding to support their breakup into smaller learning communities; funds typically were received for one planning year and two years of subsequent redesigning.
Data:	Site visited in each of the three funded years; surveyed in planning year and again two years later (in the second year of redesign). Rolling sample of schools, beginning with those that began to receive foundation funding in 2001–02 (typically their planning year).
Other schools:	Large public high schools and public high schools that impose selection criteria or offer a specialized curriculum (e.g., performing arts), located in one of the four districts selected for analysis of student characteristics and outcomes that did not receive foundation assistance.
Data:	District supplied demographic and student outcome data from 2001–02; 2002–03; and 2003–04.

- District Three: West Coast district
- District Four: East Coast district

These districts were selected because they had both multiple foundation-supported high schools and more than one intermediary working in the district. All of these districts were large, urban jurisdictions located in cities where the population ranged from 250,000 to one million people. In Districts One and Three, a majority of residents were racial/ethnic minorities. In all four districts, 20 to 30 percent of the residents live below the poverty line.

We used the districts' school-level data to measure both demographic characteristics and outcomes other than state assessment results. Demographic measures included minority composition and the percentage of students who qualify for free and reduced-priced lunch, special education services, and English language learner services. Non-assessment outcome measures included average daily attendance, progression, and suspension rates. Most of the school-level data were obtained from public district sources such as district/state Web sites. Districts provided school-level progression rates. For assessment data, we used student-level data,

allowing us to control for middle school achievement. Districts provided student-level data, including middle school and high school achievement scores and a high school identifier.

For comparison purposes, we divided high schools into three basic types:

- new high schools,
- redesigned high schools, and
- other high schools in the district.[2]

These "other" high schools included both comprehensive schools that enrolled students primarily based on the students' residence and selective or specialty schools that either offered a unique curriculum (e.g., performing arts) or were selective in admission (e.g., limited enrollment to students who had met specified middle school performance criteria).

In 2003–04, the four districts varied considerably in terms of the types of high schools in their district. Table 1 provides the counts of high schools by type in each of the four districts. In 2003–04, District One had only redesigned schools and District Four had only new schools. Districts Two and Three had both types of foundation-supported schools, with a predominance of redesigned high schools. For each of the extant data variables we examined, our analyses compared the mean value for each type (new and redesign) of foundation-supported school in a district to the average of the other high schools located in the same district.

Analyses of school demographic and non-assessment-based outcomes relied on school-level data. Analyses of assessment data were conducted at the student level. Because the number of high schools in each district was small, the statistical power of comparisons was small. Therefore, we relied primarily on the size of the difference in means between school types, rather than statistical significance tests, in the school-level analyses. Conversely, in the analyses of achievement data, the large number of students made statistical significance relatively easy to achieve, so we focused on the magnitude of differences rather than statistical significance at the student level as well.

Table 1. Number of High Schools, by School Type and District in 2003–04

District	*New*	*Redesign*	*Other*	*Total*
District One	0	7	17	24
District Two	2	11	8	21
District Three	4	7	7	18
District Four	3	0	5	8
Total	**9**	**25**	**37**	**71**

Districts measure student outcomes differently. Interstate comparisons of state assessment results, in particular, are not appropriate. For this reason, we do not present numerical values in their original metrics in our summary of findings. Rather, we have "standardized" differences by calculating effect sizes. Effect sizes are calculated by dividing the difference in means observed between groups by the pooled standard deviation. Conceptually effect sizes express differences between two groups in "units" that reflect how much individual cases vary from each other. Because our effect sizes reflect both the magnitude of the difference between foundation-supported schools and the amount of variation between schools, which is not the same in the four districts, the reader is cautioned against making direct comparisons of effect sizes across districts.

In the text, we use the labeling convention suggested by Cohen (1988a) and refer to differences with effect sizes less than 0.2 as being "similar." We denote standardized differences with an absolute value between 0.2 and 0.5 as being "slightly" higher or lower. We use "moderately" to describe effect sizes with an absolute value between 0.5 and 0.8. Finally, we labeled standardized differences that exceeded 0.8 in absolute value as being "substantially" higher or lower. For the full results of our analyses and a more detailed description of the schools and methods, please see the technical appendix.

Research Question 2: What are some of the factors (e.g., organizational characteristics of schools and teaching practices) associated with relative success on student outcomes?

To investigate the relationships between school characteristics, teaching practices, and student attitudes, we used survey data from all of the foundation-supported schools (44) surveyed in 2002, 2003, or 2004 to construct measures of these concepts. Note we are using data from our national sample schools to address Research Question 2, rather than extant data available from the four districts we used for Research Question 1. We created an "implementation index" to capture the extent to which the foundation's school attributes were present in schools.[3] We also created two separate measures of teaching practices. The reform-oriented instruction measure represents the extent to which teachers said they structured instruction to guide student-initiated research and analysis, provided for deep exploration of topics, used hands-on demonstrations and presentation, assigned group and multidisciplinary projects, used portfolios and performance-based assessments, and employed other related strategies. Conventional instructional practice represents the extent to which teachers and students take more conventional roles

in the classroom (e.g., teachers lecture to the class as a whole; lead practices on definitions, computations, and formulas; and focus instruction on preparation for standardized tests; see technical appendix for details).

We also created five measures of student attitudes:

- *Engagement-interest* — students' reports of how often they asked questions or contributed in class, met with teachers, and talked about schoolwork outside of class
- *Engagement-persistence* — students' reports of how often they gave extra effort on challenging assignments, got help with difficult homework, and resisted giving up when work was hard or not interesting
- *Academic self-concept* — the degree to which students felt they were good at reading, writing, learning mathematics, getting help, and working with others
- *Satisfaction-academic progress* — how well students felt they had been taught to read, write, analyze math problems, and learn on their own
- *Satisfaction-social responsibility* — show well students felt they had been taught to be responsible members of the community, respect diverse opinions, and think critically

When using the survey data to address the relationship among school characteristics, teaching practices, and student attitudes, we accounted for the nested nature of our survey data (i.e., students nested within schools) using hierarchical linear modeling (HLM). The HLM results presented in this report are based on 15 separate HLM models — one for each school characteristic of interest (i.e., implementation, reform-oriented teaching, and conventional teaching) multiplied by each of the five student attitude measures. Details can be found in the technical appendix.

To examine the associations among implementation, teaching, and student attitudes and student test scores, we linked school-level survey data to high school achievement data available to us for a subset of 20 schools. We restricted analysis to those 20 schools where we could readily collect 10th-grade assessment information for the sampled school and the other schools in the same district. We needed assessment data from the other schools in the district to place a given school's assessment results in proper context. State assessments have different metrics and are mapped to state-specific standards. Therefore, direct interstate comparisons are inappropriate. We relied on the comparison of each school's test scores to the performance of other schools in its district to identify relative success. We limited our analysis to sample schools where

we could ascertain the district average and distribution of assessment scores because we needed this information to facilitate analysis of results across state lines.

We supplemented both sets of quantitative analyses with qualitative data collected during site visits (see technical appendix for description of site visits and focus groups). Site visits were conducted in only a subset of the schools we surveyed. When addressing Research Question 1, we focused particularly on six of the site-visited schools that were located in one of the four districts we examined. We had our full complement of data available in these six schools: student demographic information, student-level achievement scores, teacher and student surveys, and site visits. We also included qualitative information from two additional schools for which we had survey, school-level assessment, and site-visit data. These two additional schools supplying qualitative information were selected because they had high scores on both state assessments and survey-based measures of implementation and student engagement.

To assess the relationship between student achievement and the foundation's school attributes, we used *t*-tests to evaluate the observed differences in school means. We determined whether or not each of the schools performed above or below their district's mean on the reading/English language arts (ELA) and math assessment. After identifying schools that scored above and below these district means, we used t-tests to compare the means on a series of survey-based measures: implementation, teaching practices, and student attitudes.

Results

This section describes results related to the two research questions. This first part responds to Research Question 1, beginning with an assessment of differences in student enrollment and followed by a comparison of student outcomes.

Research Question 1: How do foundation-supported schools compare with other schools in the same district in terms of the students they serve and certain student outcomes (e.g., attendance, achievement)?

Student Enrollment

Table 2 summarizes differences observed in the proportion of students from a racial/ethnic minority[4] group, receiving free or reduced-price lunch, having an individual educational plan, and being labeled an English language learner. These differences are expressed in terms of effect sizes. The closer the effect size is closer to zero, the less of a difference between the foundation-supported and other schools in that district. The greater the effect size, the larger the difference between the foundation-supported and other schools in the district. Positive numbers indicate that supported schools have higher values than other schools in their district on the measure in questions. Negative numbers indicate lower values in supported schools than in other high schools. As effect sizes were calculated separately within each district, direct comparison of effect size values across districts is not appropriate.

In all four of the districts examined, foundation-supported schools enrolled a higher proportion of students from traditionally underserved populations than did other high schools in the district. Districts Two and Three enrolled moderately to substantially higher proportions of students with these characteristics. The greater overrepresentation of historically underserved students in certain schools within these districts is likely related to the high levels of residential segregation (by race and class) there. Additionally, foundation-supported schools generally enrolled a greater percentage of students who had special educational or language acquisition needs.

In our conversations with school staff members during site visits to foundation-supported schools, we found that personnel in each of the four districts were aware of and concerned about the additional challenges many students from traditionally underserved populations face. The school personnel recognized the heightened importance of school as quite possibly the only avenue toward success for students from

Table 2. Comparisons Between Foundation-Supported High Schools and Non-Foundation-Supported High Schools in the Percentage of Students from Traditionally Underserved Populations in 2003–2004 by District (Effect Sizes in Parentheses)

District	*Non-White Minority Students*	*Free or Reduced-Priced Lunch*	*Individual Education Program*	*English Language Learner*	*Middle School Achievement Math*	*Middle School Achievement ELA/ Reading*
New vs. Other High Schools						
District Two	Substantially Higher (0.85)	Substantially Higher (1.27)	Slightly Higher (0.36)	Not Analyzed	Moderately Lower (-0.59)	Slightly Lower (-0.42)
District Three	Substantially Higher (0.83)	Moderately Higher (0.72)	Slightly Lower (-0.28)	Slightly Higher (0.47)	Slightly Lower (-0.28)	Slightly Lower (-0.38)
District Four	Slightly Higher (0.28)	Substantially Higher (1.08)	Similar (-0.012)	Moderately Higher (0.50)	Slightly Lower (-0.22)	Similar (-0.13)
Redesigned vs. Other High Schools						
District One	Slightly Higher (0.23)	Slightly Higher (0.41)	Substantially Higher (1.24)	Not Analyzed	Slightly Lower (-0.48)	Slightly Lower (-0.25)
District Two	Moderately Higher (0.71)	Substantially Higher (1.048)	Substantially Higher (0.87)	Not Analyzed	Moderately Lower (-0.63)	Moderately Lower (-0.56)
District Three	Moderately Higher (0.61)	Substantially Higher (1.91)	Substantially Higher (1.02)	Substantially Higher (1.10)	Moderately Lower (-0.63)	Moderately Lower (-0.68)

Source: Extant 2003–04 district data.

Note: Number of schools by district:
District One: Redesigned = 7; Other = 17
District Two: New = 2; Redesigned = 11; Other = 8
District Three: New = 4; Redesigned = 7; Other = 7
District Four: New = 3; Other = 6
Number of students by district:
District One: Math = 3,351; Reading/ELA = 3,404
District Two: Math = 1,821; Reading/ELA = 1,165
District Three: Math = 2,070; Reading/ELA = 1,586
District Four: Math = 1,101; Reading/ELA = 1,024
(A few schools were not able supply all of the data requested. Therefore, there was some variation in the population size by data element. See technical appendix for complete details.)

challenging home and family situations. Staff members listed a number of specific challenges their students faced, including lack of parent supervision, being parents themselves, and drugs and violence in the home. Despite staff members' willingness to work through these situations, they expressed deep concern regarding the additional needs of their students. Some school staff cited the non-native English speakers as "the most difficult" to bring up to speed to meet standards. In other schools, with lower percentages of non-native English speakers, school staff pointed to the troubled home lives of students as their biggest obstacles.

Table 2 also compares the middle school performance level of students attending foundation-supported and non-supported schools. In general, we found that students entering foundation-supported high schools were starting their secondary educations behind students enrolling elsewhere in terms of reading and math achievement. In only one of the eight comparisons (District Four, ELA), were students' prior achievement levels similar to those of students attending other high schools in the district.

During site visits in these districts, school staff in foundation-supported high schools expressed consternation regarding students arriving "without the necessary basic skills" and with academic foundations that were "so weak." Although concerns about math skills were often voiced as well, staff members we spoke with across all four districts consistently identified literacy as being the weakest link in students' skill sets. They noted that lessons and units could take twice as much time to complete because, as one teacher stated, "writing, spelling, reading, and grammar aren't there with most of the students." Another teacher described their students as being "as many as five years below grade in math and/or reading" upon entering high school.

Student Outcomes: Achievement

This disparity leads to critical questions about differences in student outcomes, as foundation-supported schools and their teachers work with students who enter with greater challenges to overcome. Table 3 summarizes findings related to the second part of this research question—differences in student outcomes. While foundation-supported high schools should be held to the highest standards, it is important to acknowledge that they often serve students who begin secondary education academically behind their peers. Therefore, determining the effectiveness of new and redesigned schools needs to incorporate a "value added" component in addition to unadjusted outcomes. The first two columns of table 3 provide simple comparisons of test scores in math and ELA, without controlling for prior achievement. These columns suggest that new schools scored slightly lower on ELA than other schools in their districts. The math achievement scores for new schools varied, ranging from similar to moderately lower. Where assessment data were available, redesigned schools performed moderately lower in both of these content areas compared to other schools in their districts. After controlling for prior achievement (right column), these differences tended to narrow but did not disappear.

Table 3 provides a snapshot for the 2003–04 academic year, comparing foundation-supported schools with other schools in their districts. Such comparisons provide limited information about the status of foundation-supported schools and should be interpreted with caution. The foundation-

Table 3. Comparisons of High School Achievement Test Scores by District (Effect Sizes in Parentheses)

District	*Simple Comparisons of High School Assessment Math Scores*	*Simple Comparisons of High School Assessment English Language Arts/ Reading Scores*	*Effect of Controlling for Middle School Achievement*
New vs. Other High Schools			
District Two	Moderately Lower (-0.74)	Slightly Lower (-0.28)	Math disparity reduced, ELA disparity eliminated
District Three	Slightly Lower (-0.46)	Slightly Lower (-0.24)	Both math and ELA disparities remain
District Four	Similar (-0.18)	Slightly Lower (-0.24)	Math scores remain similar, ELA disparity eliminated
Redesigned vs. Other High Schools			
District One	Data not available	Moderately Lower (-0.54)	Math data not available, ELA disparity reduced
District Two	Moderately Lower (-0.68)	Moderately Lower (-0.56)	Both math and ELA disparities reduced
District Three	Moderately Lower (-0.75)	Moderately Lower (-0.66)	Both math and ELA disparities reduced

Source: Extant 2003–04 district data.

Note: Number of schools by district
District One: Redesigned = 7; Other = 16
District Two: New = 2; Redesigned = 9; Other = 8
District Three: New = 3; Redesigned = 5; Other = 3
District Four: New = 3; Other = 7
Number of students by district
District One: Reading/ELA = 3,043
District Two: Math = 1,813; Reading/ELA = 1,810
District Three: Math = 2,055; Reading/ELA = 2,032
District Four: Math = 1,765; Reading/ELA = 1,765

supported schools enrolled higher concentrations of lower achieving (and therefore more challenging) students than other schools in their districts and were in either their initial year or year two of their intervention.

Better indicators of their success are an examination of improvement, or progress achieved as they implemented the foundation's initiative, and comparison of these trends with those of other schools in their districts. To date, we are able to provide only a limited analysis of trends in achievement because many of the schools were too new to have two years of data, and there is significant lag time in the availability of such extant data. Because high school assessments are typically administered in the 10th grade, a new school usually needs to be open for four years before a trend in achievement can be monitored, assuming the need to wait until several months into a school's fourth year for the achievement data from year three to be released.

Table 4 provides trend data for three of the four districts. (Trend data for District 3 are not available because a new assessment instrument was introduced in 2003–04.) As illustrated in the table, some evidence suggests that some foundation-supported schools may be closing the gap in reading and ELA scores. In Districts One and Four, the increase in test scores from 2003 to 2004 was higher in foundation-supported schools than in other schools in their jurisdictions. Math results were mixed — a smaller increase for foundation-supported schools in one district, a larger increase in another.

The nature of the trend results for District Four are different from those for Districts One and Two, which compare the performance of two sequential 10th-grade cohorts. District Four changed the timing of its high school assessment tests from testing in 10th grade in 2002–03 to testing in 11th grade in 2003–04. Therefore, the year-to-year differences in District Four are for the same cohort of students, the class of 2005. District Four's class of 2005 who were enrolled in new schools made substantially higher gains in their reading and ELA skills than students enrolled in non-foundation-supported schools between their sophomore and junior years. But these same students experienced a moderately smaller gain in math skills than did students in other district schools during this interval.

As previously mentioned, during our site visits, school staff described students in these foundation-supported schools as entering with below-grade-level content mastery in almost all subjects. The schools we visited were unanimously focused on literacy as their top academic priority. Schools had instituted various methods to strengthen students' reading/ELA scores on state assessments. These methods included adopting

Table 4. Trends in High School Achievement Test Scores by School Type and District

District	*Amount of Increase in 10th-Grade Math Scores 2003 to 2004*	*Amount of Increase in 10th-Grade ELA/Reading Scores 2003 to 2004*
New vs. Other High Schools		
District Four	Moderately Lower (-0.54)	Substantially Higher (1.44)
Redesigned vs. Other High Schools		
District One	Data not available from district	Slightly Higher (0.28)
District Two	Moderately Higher (0.79)	Similar (0.01)

Source: Extant 2002-03 and 2003-04 district data.
Note: Number of schools by district:
District One: Redesigned = 4; Other = 15
District Two: New = 2; Redesigned = 5; Other = 8
District Four: New = 3; Other = 7

pre-existing reading programs, such as AVID and High Point, implementing "reading workshops" in class that allow for a minimum of 10 minutes of students' reading, and having students adjust to "drafting" writing assignments so that they are "just kind of constantly looking at the writing," according to teachers. In the eyes of the teachers we talked to, these various methods seem to be working, as the teachers already describe project work as getting "better" and "deeper" and student writing getting "much, much better." The following, from an English teacher who looped with his 9th-grade students to 10th grade, is an example of the perceived improvement with reading and ELA:

> "I think there has been some growth in kids' writing, especially since they had me again, so the language is the same, and that has been helpful. So they know things like 'hook,' 'background,' 'thesis,' 'evidence,' and 'conclusion.' The kids took the high school exit exam two weeks ago, as sort of a trial run. There's a writing piece on that, and honestly, I didn't do explicit prep for it. When they got to the essay, they said, 'Mr. Smith, this is an essay. Do we need to do a hook and a background?' and that was gratifying that they've internalized that. To a degree, that's progress."

The staff members with whom we met described fewer specific methods geared toward improving math test scores, although one teacher noted that "the culture with respect to attitudes toward math has changed among the students, but this year's scores won't reflect that." Another school's teacher noted the loss of two math teachers during the academic year and the school's inability to find long-term replacements as a source of worry that "the results on the (math) test for these students will be terrible."

Student Outcomes: Behaviors

In addition to state assessment outcomes, we examined average attendance rates, progression rates, and suspension rates for high schools in these four districts. Attendance rates provide an important first glimpse of high schools' early success in changing school culture concerning student behavior. Progression rates provide an estimate of the percentage of entering 9th-grade students who graduate within four years. (Progression rates are calculated by multiplying the percentage of students who progress from grade 9 to grade 10, from grade 10 to grade 11, from grade 11 to grade 12, and from grade 12 to graduation at the expected pace, resulting in a degree in four years.) The number of suspensions per 100 students served as a proxy measure for problematic student behavior. Although a high number of suspensions may indicate a school is cracking down on problematic students, high-performing schools usually do

not have high suspension rates. It is important to note that the suspension rate is not the same as the percentage of students who have been suspended; a single student can be suspended multiple times. Table 5 summarizes the data on these three measures.

We found mixed results for student behavior measures in the foundation-supported schools compared with other high schools in the same district. As a group, new high schools enjoyed higher attendance rates than the other schools in both districts that had attendance data available. We saw the opposite pattern for redesigned schools. In both Districts One and Two, redesigned schools had lower attendance rates than the other schools in the district. The comparisons of progression rates produced mixed results for new schools, but results were consistently lower in redesigned schools compared to other schools in their districts. These two findings for the two foundation-supported school types do not mean that new schools are necessarily more successful in promoting positive student behavior. The decision to redesign a high school is based in part on poor performance in the past. Therefore, many redesigned high schools are dealing with a legacy of problematic student behavior. There was no discernable pattern to our finding for suspension

Table 5. Comparisons of Student Behavior in Foundation-Supported High Schools and Non-Foundation-Supported High Schools in 2003–04 by District and Type of School

District	*Attendance*	*Progression Rate*	*Suspensions per 100 Students*
New Schools vs. Other High Schools			
District Two	Moderately Higher (0.74)	Substantially Lower (-2.22)	Moderately Higher (0.45)
District Four	Moderately Higher (0.68)	Substantially Higher (1.99)	Similar (-0.01)
Redesigned Schools vs. Other High Schools			
District One	Substantially Lower (-1.19)	Substantially Lower (-1.09)	Similar (0.11)
District Two	Slightly Lower (-0.31)	Substantially Lower (-1.57)	Moderately Higher (0.54)

Source: Extant 2003–04 district data.
Note: Number of schools by type and district:
District One: Redesigned = 7; Other = 17
District Two: New = 2; Redesigned = 11; Other = 8
District Four: New = 3; Other = 6
(A few schools were not able supply all of the data requested. Therefore, some variation in the population size occurred by data element. See technical appendix for complete details.)

rates. Suspension rates were similar to the district norm among the new schools of District Four and the redesigned schools of District One, but higher among both new and redesigned schools in District Two.

Summary

In addressing our first research question, we compared foundation-supported schools to the other schools located in the same district. Based on comparisons of student demographic and prior achievement data, we conclude that grantees working in the four districts examined have been initially successful in funding schools that enroll underserved populations of students. Based on comparisons of state assessment data in middle school, we conclude that, on average, students who attended foundation-supported schools during the 2003–04 academic year entered high school not performing as well as students enrolled in other high schools in their districts. Thus, the schools started under the foundation's initiative are on the whole serving students who enter high school with critical challenges. Based on the limited amount of data available thus far, there is some initial evidence that foundation-supported schools are making progress in narrowing this achievement gap (between foundation-supported and other schools in the district) for reading/ELA. On the other hand, we found little evidence of a narrowing of the achievement disparities in mathematics. Our comparisons of student behavior measures produced mixed results. We found that new schools had higher attendance rates, but redesigned schools had lower attendance than did the rest of the schools in the same district.

Research Question 2: What are some of the factors (e.g., organizational characteristics of schools and teaching practices) associated with relative success on student outcomes?

As we turn to our second research question, we stop comparing foundation-supported schools to the other high schools in the same district. Instead, we make comparisons among foundation-supported schools. We examine the strength of the relationships that connect the components of the foundation's strategy for improving secondary education: school attributes are expected to affect classroom teaching which affects student attitudes which affect student outcomes. These comparisons use the responses teachers and students provided on surveys administered by AIR®/SRI to assess the strength of the associations among school attributes, teaching practices, and student attitudes as well as the relationship between each of these school qualities and student achievement. As we have readily acknowledged, the cross-sectional data we have available are not sufficient to determine causality, but our data can determine which of the presumed associations are observed in reforming high

schools. We used the foundation's sequential strategy for improving high schools to organize our presentation of results. Specifically we addressed five sets of associations:

- Relationships between implementation of school attributes and student attitudes;
- Relationships between teaching practices (reform-oriented and conventional) and student attitudes;
- Relationships between implementation of school attributes and state assessments of academic achievement;
- Relationships between teaching practices and state assessments of academic achievement; and
- Relationships between student attitudes and performance on state assessments of academic achievement.

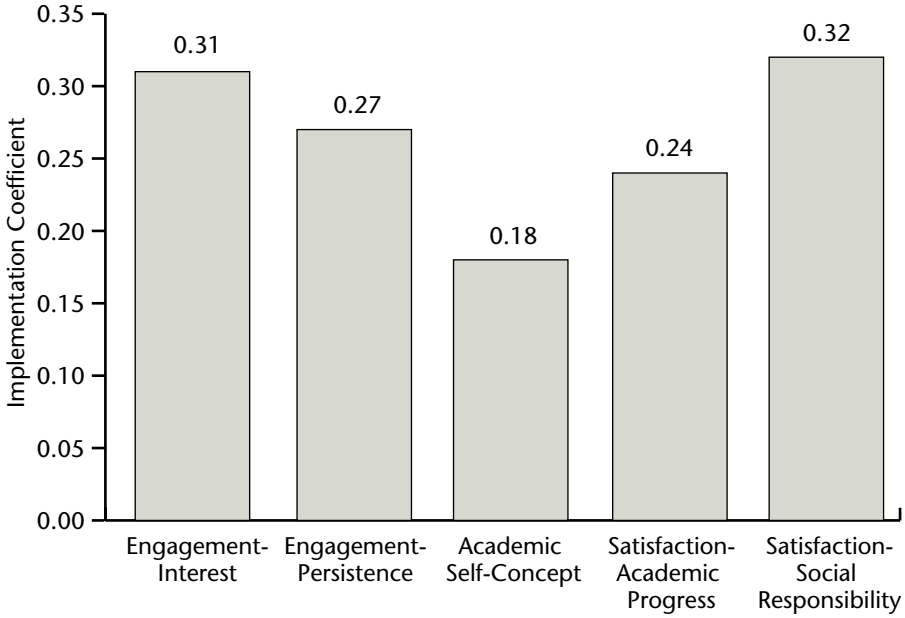
Note that we do not address the relationships between implementation and teaching practices because we explored the association between these two components in great detail in our earlier report, *Rigor, Relevance, and Results in New and Conventional High Schools* (2005). That report examined how the implementation of desired school attributes helped foster reform-oriented teaching practices over time.

Student Attitudes

School Attributes and Student Attitudes. The literature suggests that students must actively engage with their education before learning can occur. Hence, positive student attitudes are a presumed precursor to improved student performance. Are the school attributes championed by the foundation associated with better student attitudes? To answer this question, we used the survey-derived implementation index and five student attitude measures reflecting students' opinions about themselves as students, their schoolwork, their teachers, and satisfaction with the education they were receiving (see technical appendix for details). Figure 1 presents these relationships and indicates that implementation of the foundation's school attributes was positively related to all five student attitudes ($p < .01$). In schools where a higher degree of the foundation's attributes of effective schools were implemented, we found students more engaged, more persistent in their learning, and more satisfied with their academic progress. These results control for student background characteristics, a school risk index, and the year of survey administration.

We found strong corroborating evidence in the schools where we conducted site visits. The schools we visited promoted sustained relationships

Figure 1. Relationship Between Student Attitudes and Implementation



Source: The figure represents coefficients from our HLM analysis. Data used in the model were from the school information form, teacher and student surveys from 2001–02, 2002–03 and 2003–04. Detailed output is provided in the technical appendix.

Note: Number of schools: 44; number of students: 10,887.

between and among adults and students, based relationships on mutual respect and responsibility, and had staff with high expectations. We found students who were engaged and persistent in their learning. Students also took their schoolwork seriously, found it challenging, and were confident in their accomplishments. One specific illustration of the link between school attributes and student attitude is provided by a 10th-grade student who said he had once been negative about schools but explained that the mutual respect among students and teachers at the school, the greater personal responsibility required at the school, and the inherent trust that governs the school's community were reasons for his re-engagement with school. Confirming these sentiments, students in another foundation-supported school's focus group collectively noted that they have different perceptions of their schoolwork now, compared to other schools, because they were actually "passionate" about what they are doing and felt that they appreciated the value of their education more than they had when they were attending regular high schools. According to one student,

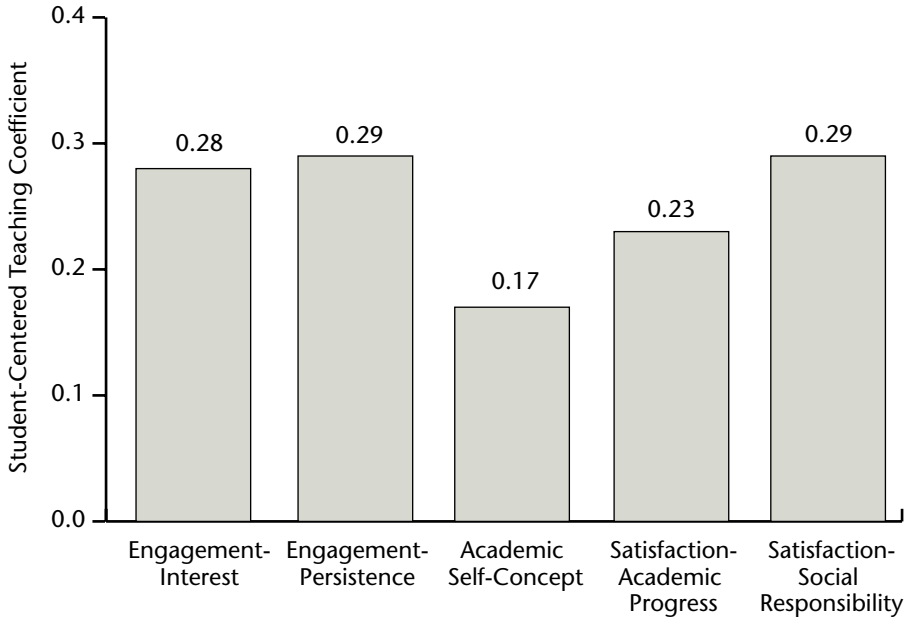
> "Most kids are like, 'well I know I'm not going to be using geometry on my job,' but they don't realize you're going to be

using math every day. [Those students] drop out of class, and then they learn later that they're going to need to use it. Here, we can see it every day."

A teacher at the same school discussed the notion that students were making meaningful progress towards their learning goals and honoring their interests: "Kids who were apathetic have become excited again. Kids aren't tuning out. At other schools, kids stop being interested at high school."

Teaching Practices and Student Attitudes. Our framework suggests that, in addition to school attributes, effective teaching is an important factor influencing student attitudes and outcomes. As described in the methods section, we developed two measures of teaching practices: reform-oriented teaching and conventional teaching. Our reform-oriented measure captured the degree to which teaching called for students to be active learners and for teachers to act as guides. Reform-oriented teaching also involves in-depth learning, portfolios of student work, and using technology as a tool. Conventional teaching, on the other hand, reflects an emphasis on rote memorization of facts, lectures to the class as a whole, and the teacher taking the lead on discussions.

Figure 2. Effect of Reform-oriented Teaching on Student Attitude



Source: The figure represents coefficients from our HLM analysis. Data used in the model were from the school information form, teacher and student surveys from 2001–02, 2002–03 and 2003–04. Detailed output is provided in the technical appendix.

Note: Number of schools: 44; number of students: 10,887.

Figure 2 illustrates the relationship between reform-oriented teaching and student attitudes. Students had more positive attitudes when a greater degree of reform-oriented instruction was used in the classroom. Students were more satisfied with instruction on social responsibility matters and were more persistent and engaged in their learning in classrooms where more elements of reform-oriented teaching were in place. Consistent with the findings for implementation, all relationships with the five student attitudes were statistically significant.

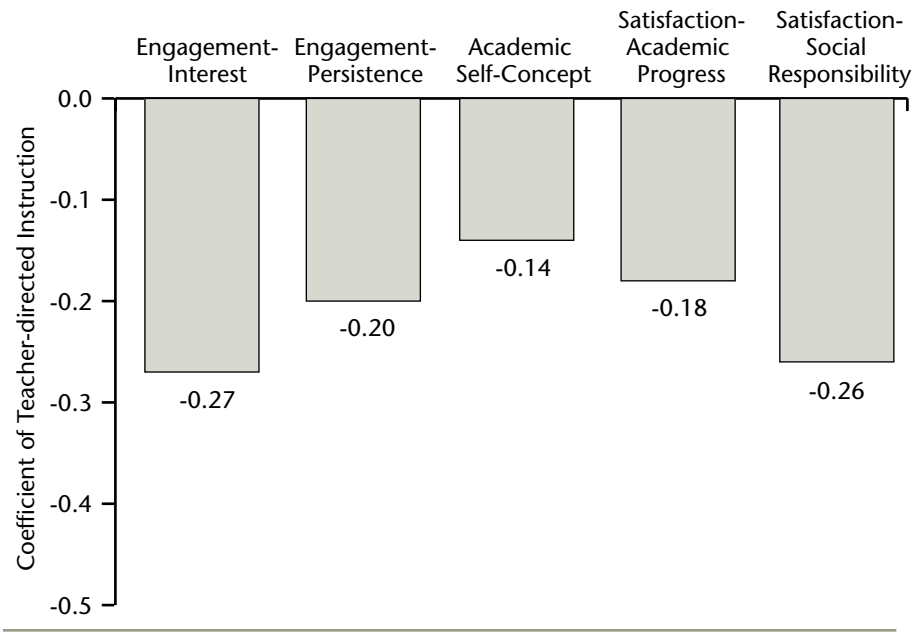
The students we spoke with during focus groups provided insight into the nature of the relationship between reform-oriented instruction and student attitudes. Students shared their opinions that in-depth learning and student involvement in academic content selection made them more interested in their learning and increased their academic self-concept. One student explained that students are allowed to choose their projects based on their own interests; this choice fostered creativity and prompted her to work harder: "When you want to learn about something, you get a lot out of [the schoolwork]. When you want to learn about [subject] it's more fun. I'm going to spend more time on it because it's fun."

She described two of her projects, one on the history of soccer and the other on appendicitis, both of which had personal relevance for her. Project work based on students' interests provided students with an enhanced sense of accomplishment. Another student illustrated this by saying,

> "We always have something like a final product, which before you really don't have that at a high school. You do this one assignment and then it will be gone forever—you have no idea where it went, what it was, anything. And then here it's like everything always comes together for a final product, like we made something. So it feels like we accomplished—I feel like I have accomplished more because you see this thing that you did, not just a bunch of papers laid out."

In contrast, conventional instruction, as illustrated in Figure 3, was negatively related to all student attitudes. The strength of these negative relationships with the individual student attitudes closely mirrored the corollary positive relationship between each student attitude and reform-oriented instruction. As these results are based on cross-sectional correlations, we can not be sure of the direction of causality. It is possible that reform-oriented teaching practices are actually improving student attitudes, but it is also possible those positive student attitudes enable reform-oriented instruction or that students with positive attitudes were more likely to choose schools where reform-oriented instruction was offered. It seems likely that the causality within this teacher-student dynamic runs in both directions.

Figure 3. Effect of Conventional Instruction on Student Attitudes



Source: The figure represents coefficients from our HLM analysis. Data used in the model were from the school information form, teacher and student surveys from 2001–02, 2002–03 and 2003–04. Detailed output is provided in the technical appendix.

Note: Number of schools: 44; number of students: 10,887.

Next, we examine the relationship among school attributes, teaching practices, and student attitudes with the student academic achievement as measured by state assessments.

Student Achievement

Our analysis of student achievement combines the results from schools in a number of states. The variation in state assessments presented an analytical challenge. The high school achievement tests each state used were unique. Each state's assessments were mapped to different educational goals and were measured by instruments tailored to measure progress toward those goals. States also varied in terms of the timetable they have set for themselves in achieving student proficiency standards based on No Child Left Behind (NCLB), so interstate comparisons of the percentage of students passing exams are not appropriate.

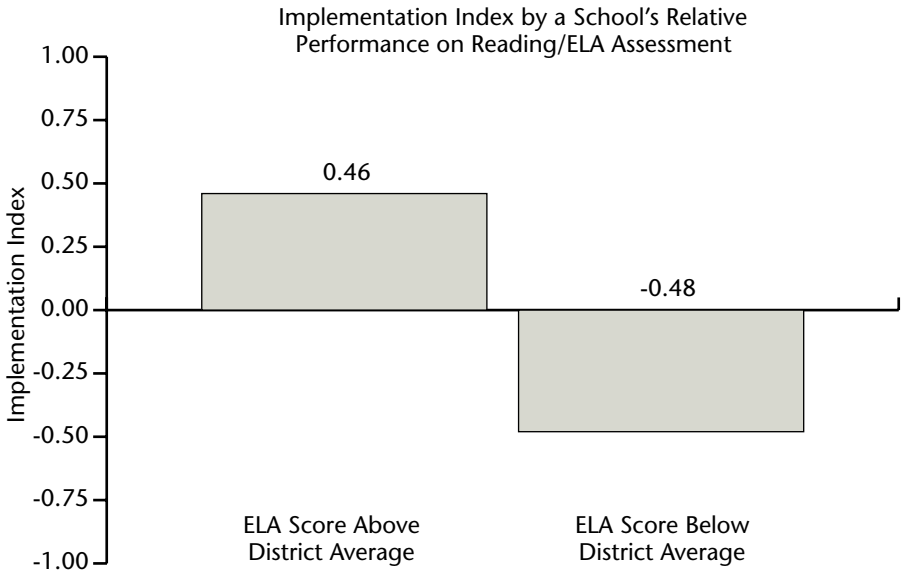
Lacking a common metric to compare assessment results, we divided foundation-supported schools into groups based on their performance on state reading/language arts and math assessments relative to the performance of other schools in their districts. As previously explained in the study-design section, we limited our analysis of achievement results to the subset of 20 surveyed schools where school and district achievement

data were both readily available. We divided the foundation-supported schools with survey and achievement data into two groups: (1) schools that scored above the district mean in reading/ELA, and (2) schools that scored below the district mean in reading/ELA. We created similar dichotomies in mathematics. We then compared the mean level of school attributes, teaching practices, and student attitudes in the groups defined by the school's relative performance on state assessments.

Student Achievement and Implementation. When we compared the school attributes across groups defined by assessment results, we found the mean value of our implementation index was significantly higher in schools that scored above the district mean on reading/ELA tests than in those that scored below the district mean (see Figure 4). We did not find a significant difference in implementation level between the two achievement groups defined by math.

Perhaps higher levels of personalization, expectations, respect and responsibility, common focus, collaboration, and using technology as a tool contributed to higher levels of student performance on reading/ELA tests. As our data are cross-sectional, the result might also be due to schools serving students with lower literacy levels (i.e., schools scoring below the district mean) having a harder time implementing the foundation's school attributes. Likewise, the lack of a relationship between school attributes and math scores has two plausible explanations — either

Figure 4. Student Achievement and Implementation



Source: Measures of implementation were garnered from survey data. School ELA scores were aggregated from district-provided assessment scores.

Note: Total number of schools = 19 (12 above district mean and seven below district mean). See technical appendix for detailed results.

Table 6. Relationship Between Student Outcomes and Teaching Practices

Subject	*Average Reform-oriented Teaching Scale*	*Average Conventional Teaching Scale*
English Language Arts		
Schools Scoring Above District Mean	.31	-.23
Schools Scoring Below District Mean	-.17	.43
Mathematics		
Schools Scoring Above District Mean	-.23	.07
Schools Scoring Below District Mean	.29	.02

Source: Measures of teaching practices were garnered from survey data. School ELA and mathematics scores were aggregated from district-provided assessment scores.

Note: Total number of schools = 19 (12 above district mean, seven below district mean). See technical appendix for detailed results.

student math skills may not be influenced by high school attributes or the lack of student math skills may not pose a barrier to schools' implementation efforts. As we turn to relationships among student achievement and teaching practices and student attitudes, we will see this same pattern of positive associations involving reading/language arts but no association or negative associations involving math.

Teaching Practices and Student Achievement. After finding that teachers' reports of their instructional practices were related to student-reported attitudes, we looked for relationships between teaching practices and student achievement test scores. Although none of the relationships between teaching practice and student achievement measures attained statistical significance, measures of reform-oriented teaching were positively associated with achievement in ELA and negatively associated with achievement in mathematics. The inverse was found for conventional instruction. Conventional instruction was negatively associated with achievement in ELA. The amount of conventional instruction at schools scoring above and below the district mean in mathematics was quite similar. The relatively small number of schools in these analyses contributed to the lack of statistically significant findings.

These data suggest that reform-oriented instructional techniques either promote or require ELA skills. At the same time, reform-oriented instruction was negatively associated with the level of math skills. Either reform-oriented instruction failed to foster or did not require a high level of student proficiency in mathematics. The challenge that mathematics instruction poses was addressed in AIR/SRI's report, *Rigor, Relevance, and Results in New and Conventional High Schools* (2005). As reported there, many faculty members were not wholly satisfied with their instructional programs in mathematics. Teachers described many barriers to rigorous mathematics, including students' weak numeracy skills, the limited

availability of qualified math teachers, the paucity of useful instructional resources and professional development offerings, and the difficulty of integrating mathematics with other content.

Excerpts from this earlier report indicate how teachers talk about these challenges. One teacher described her response to students' low numeracy skills, saying, "So in some sense—I don't know if 'regressed' is the right word—but I've backed away from some of my big expectations."

Describing her school's difficulty developing math curricula, another teacher said,

> "There isn't a big market for materials for this kind of learning [project-based learning in math] in this kind of environment. There just isn't a lot of cooperative learning going on in math. There is a lot of conflict in the field about the direction in which math curriculum should go."

Finally, another teacher summarized the difficulty integrating math with other course content,

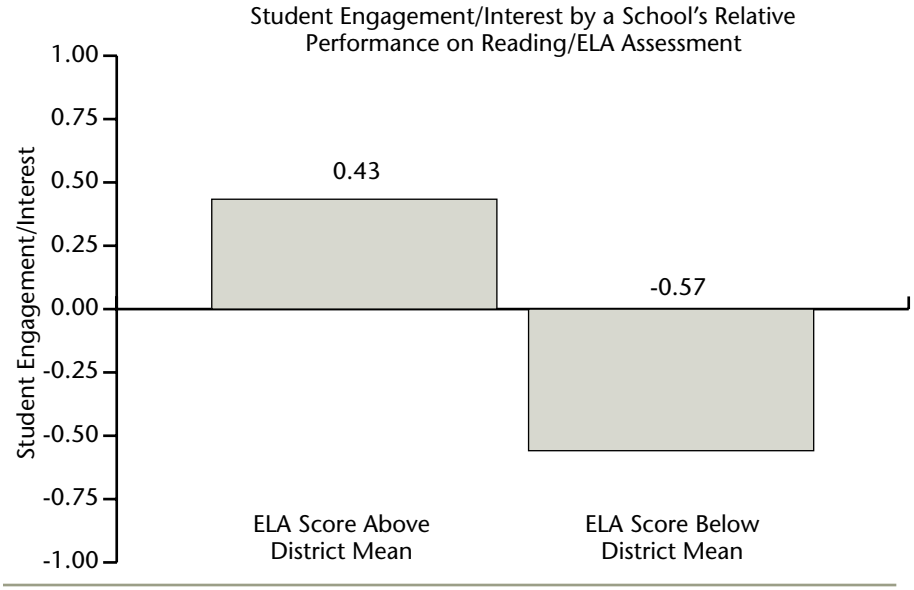
> "It's hard to find math people who look at literacy across the curriculum.... It's about being able to express the math concepts in the context of what they are writing. What does it mean to talk about the percentage of the oil that is spilled? Can you express it in the context of a paper? It's hard to find someone who can teach kids how to do that."

The relative difficulty in reforming instructional practice in mathematics may partially account for the lack of a substantive difference in the level of conventional instruction in schools scoring above and below the district average on math. Although not significant, there is a negative association between conventional instruction and performance on reading/ELA assessments. Conventional instructional techniques seem either to retard the development of reading/ELA skill or to be more likely to be used by teachers when students have low skill levels. The temptation to fall back on traditional conventional methods may become quite strong for high schools that are reforming and face concerns about their test performance.

Student Attitudes and State Assessment Results. Prior research has found a relationship between student attitudes and achievement. Of the five student attitude measures we used, only student engagement/interest was significantly related to reading/ELA test scores (see Figure 5). Both the statistically significant and non-significant findings are interesting.

Our "engagement-interest" measure reflects the frequency of behaviors indicating interest in education (e.g., contributing in class, talking about schoolwork outside of class, etc.). It is perhaps not surprising that this

Figure 5. Student Achievement and Student Engagement



Source: Measures of student engagement/interest were garnered from survey data. School ELA scores were aggregated from district-provided test scores.

Note: Total number of schools = 20 (13 above district mean, seven below district mean).

of all the student attitude measures is the one most positively related to ELA performance. Recall that many foundation-supported high schools focused their curriculum on fostering literacy skills. The level of student interest in such schools could positively influence the amount of learning that occurs; alternatively, the level of interest students have in a literacy-focused curriculum could be influenced by prior achievement levels. A school staff member shared an illustration of how her students displayed interest in their education by the approach they take to standardized tests. This teacher explained that even though some students missed passing the state exit exam by a few points, all of the students were eager to know their test scores. This teacher described students seeking out additional assessments of the academic skills. She reported not only more students requesting PSAT fee waivers this year in comparison to past years, but also more students asking teachers for assistance on specific sections on assessments, such as analogies. She pointed out that some students have come to her for help on Saturdays to prepare for testing.

None of the other student attitudes items exhibited statistically significant relationships with either ELA or math assessment scores. Although these other results were not significant, they again suggest very different implications for ELA and math instruction. There was virtually no difference in how well students felt they had been taught to read, write, and so forth among schools that scored above and below the district mean on reading/ELA assessments, but the other three student attitudes measures

were all positively associated with performance ELA tests. In contrast, all five student attitude measures were negatively related to math performance. That is, student attitudes were more positive in schools that scored below the math mean than in schools that scored above the average math score.

Summary

Our examination of the relationships involving student attitudes, school attributes, and teaching practices produced results consistent with the implied strategy for improving high schools. We found more positive student attitudes across all five attitudinal dimensions examined in schools that had implemented more of the school attributes championed by the foundation. We also found a positive association with all five of our student attitudes measures and reform-oriented instruction. Finally, we found lower levels of all measured student attitudes in schools that employed a larger degree of conventional instruction.

Our findings for the relationships with the level of academic achievement have mixed implications. We found a statistically significant association between performance on reading/ELA assessments and the implementation index. There was also a significant positive relationship between reading/ELA performance and the level of student interest. Although not statistically significant, we also found other positive reading/ELA associations between reform-oriented instruction as well as associations with most of the other student attitude measures. We did not find a single significant association among the factors we examined and math achievement. Not only did we fail to find any evidence of a positive association between foundation-supported school attributes and either teaching practices or student attitudes, the direction of most of these relationships was negative. That is, schools scoring below their district means had higher levels of the desired school factor.

Implications for the Initiative

Our first evaluation report, *Creating Cultures for Learning: Supportive Relationships in New and Redesigned High Schools* (2005), documented the progress of reform in new small and redesigned high schools, giving special attention to the development of relationships between and among students and teachers. Our second report, *Rigor, Relevance, and Results in New and Conventional High Schools* (2005), found that school-wide changes in foundation-supported sites have, indeed, reached into the classroom. This report focuses on early outcomes for students, both in terms of student attitudes, as expressed on student surveys and focus groups, and in terms of school-level performance on state assessments.

Findings

Both of the preceding reports described general progress during the first three years of the foundation's initiative; however, they found neither uniform nor linear progress in terms of schools implementing desired school attributes and instilling reform-oriented teaching practices. Given the variation in progress foundation-supported schools experienced in implementing desired school attributes and implementing reform-oriented instruction techniques, it is not surprising that the student outcome findings observed in foundation-supported schools at various stages of development are also mixed. We have found that:

- ***Foundation-supported schools are enrolling traditionally underserved populations.*** In all four of the districts examined, foundation-supported schools enrolled a higher proportion of students who were eligible for free or reduced-priced lunch and who were members of a race/ethnic minority group than did other high schools in the district. Foundation-supported schools generally enrolled a greater percentage of students with special education or language acquisition needs as well. Additionally, foundation-supported schools generally enrolled students who began high school academically behind students attending other schools in the same district.

- ***Patterns in student-level achievement data suggest promise for reading/ELA achievement, but not for mathematics.*** In two of the three districts where trends on state assessment data could be examined, we saw larger improvements in reading/ELA achievement over time in foundation-supported schools than elsewhere in the district. The third district experienced improvements to reading/ELA on a par with

the rest of the district. The size of the gain in reading was larger in the district with new high schools than it was in the district with redesigned high schools. The trend over time was mixed for math.

- ***Attendance rates were higher at new schools, but lower in redesigned schools, compared to other schools in the same district.*** As a group, new high schools enjoyed higher attendance rates than the other schools in both of the districts with attendance data available. Redesigned schools, on the other hand, had poorer attendance than the other schools in their districts.

- ***Implementation and reform-oriented instruction were positively related to student attitudes.*** In schools with higher levels of the foundation's attributes of effective schools in place, we found students more engaged, more persistent in their learning, and more satisfied with their academic progress. Students had more positive attitudes also when a greater degree of reform-oriented instruction was used in the classroom. The higher the level of conventional instruction in the classroom, the less positive student attitudes were. These analyses controlled for student background characteristics, a school risk index, and the year of survey administration.

- ***Implementation and student engagement interest were positively related to school-level English language arts performance.*** We found a significantly higher level of implementation of the foundation's school attributes in those foundation-supported high schools that scored above their district's mean on reading/ELA tests than in those that scored below the district mean.

- ***Implementation and student engagement interest were not positively related to school-level mathematics performance.*** In fact, all the associations we observed were negative. That is, implementation levels, reform instruction, and positive student attitudes tend to be higher in foundation-supported high schools that score *below* their district mean on math assessments than in schools that scored above their district's mean. None of these negative relationships attained statistical significance, however.

Implications

These findings have several implications:

- ***Judgments of secondary schools' performance should take into account the differential levels of academic and attitudinal preparation students bring to high school.*** Foundation-supported high schools are successfully enrolling the populations targeted by the initiative, and

in doing so they are enrolling high proportions of high-need, low-achieving students. Enrolling and retaining high proportions of these historically underserved students often introduce challenges which affect a school's ability to obtain strong student outcomes. Research on classrooms suggests that teachers of classes with large numbers of low-achieving students perceive constraints on the instructional content and techniques they can use (e.g., the need to focus more on basics and discipline and less on deep content and student initiative). Similarly, research on schools suggests that high concentrations of low-achieving students are associated with problems such as higher levels of delinquency (for a review of the literature, see Lee et al., 1999). Our own evaluation work describes many of the challenges teachers in foundation-supported schools report as they work to educate high-need students (AIR/SRI, 2004). Schools with high concentrations of high-need students must work extremely hard to obtain outcomes comparable to those of schools with low-need students. Additionally, to the extent that foundation-supported schools enroll more high-need students and are more successful than other schools in retaining these students, their test scores and other outcome data will be depressed. The foundation may want to focus as much as possible on "value-added" definitions of school performance. While it should never allowed to be used as an excuse, the low level of academic skills with which many students enter foundation-support high schools needs to be kept in mind. High schools should be given credit for the gains in academic mastery their students demonstrate, even if what they learn should have been mastered in middle or even elementary school. At the school-design stage, there is a need for decision tools to help design teams think carefully about the match between their organizational capacity and the needs and aspirations of the students they are recruiting.

- ***Schools need more support around mathematics content and instruction.*** In this report, we found no evidence of foundation-supported schools making progress in mathematics achievement. In part, this lack of progress may reflect some foundation-supported schools' stress on individualized learning programs built around student interests that sometimes treat mathematics in a cursory manner. Other potential contributing factors have been described in earlier reports: Some foundation-supported schools have struggled with hiring and retaining qualified math teachers, and several schools have had to build programs with teachers who were ill-prepared to teach math. Teachers have reported that good mathematics curricular materials consistent with their instructional philosophies are

hard to find, and multidisciplinary resources are particularly elusive *Creating Cultures for Learning: Supportive Relationships in New and Redesigned High Schools* (AIR/SRI, 2005). Moreover, we have found wide variation in the rigor and relevance of mathematics assignments in foundation-supported schools. On average, the rigor of math assignments in foundation-supported schools was not significantly better than that of math assignments in large, comprehensive high schools *Rigor, Relevance, and Results in New and Conventional High Schools* (AIR/SRI, 2005). Curriculum materials, professional development, technical assistance, and coaching around mathematics content and instruction are all very much needed in these schools, as are well-qualified math teachers.

- ***It is too early in the initiative to draw definitive conclusions concerning student outcomes.*** The schools included in these analyses have been serving students for three years or less. We concentrated our analysis on only four urban districts that had a relatively large number of support-schools. Patterns may well change as the initiative and evaluation expand to additional districts, additional years' worth of data become available, and schools continue their reform efforts. School-level outcome data can be very unstable from year to year, especially in schools with small enrollments. Prior research on school improvement efforts suggests that a time frame of five to six years is appropriate for assessing the viability of an educational intervention.

Working with state education agencies, school districts, and other educational organizations, the foundation has been successful in reaching the students most in need of improved secondary education. The replication of the positive relationships reported by the guiding literature involving school attributes, classroom instruction, and student attitudes indicates that reform efforts are touching the lives of young people. Our ability to gauge the impact of reform efforts on specific student outcomes such as state assessment tests and behavioral measures is, however, limited at this time. It is limited by the number of schools for which we have both measures of implementation and outcome data over a period of time. Many of the foundation-supported schools analyzed here are in their first and second year of existence. We collected information during the 2004–05 school year from these schools as they moved into their second and third years. We also began a new wave of schools that opened during 2004–05. As AIR®/SRI continue to collect implementation information from subsequent "waves" of supported high schools and to gather outcome data over time for existing cohorts of schools, we will be increasingly able to address these issues.

References

American Diploma Project (2004). *Ready or not: Creating a high school diploma that counts*. Washington, DC: Achieve, Inc.

American Institutes for Research/SRI International. (2003). *Charting a course: Evaluation design of the National School District and Network Grants Program.* Paper presented at the 2003 Annual Meeting of the American Educational Research Association, Chicago, IL.

American Institutes for Research/SRI International. (2005). *Creating cultures for learnings: Supportive relationships in new and redesigned high schools.* Menlo Park, CA: SRI International.

American Institutes for Research/SRI International. (2005). *Rigor, relevance, and results in new and conventional high schools.* Menlo Park, CA: SRI International.

Bill & Melinda Gates Foundation. (n.d.). *Helping all students achieve* [Pamphlet]. Seattle: Author.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (eds.). (1999). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Brophy, J. E. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology, 75*(5): 631–661.

Cohen, J. (1988a). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cohen, M. (1988*b*). *Restructuring the education system: Agenda for the 1990s.* Washington, DC: National Governors' Association.

Conley, D. T. (1993). Restructuring: In search of a definition. *Principal, 72*(3): 12–16.

Connell, J. P., & Kubisch, A. C. (1998). Applying a theory of change approach to the evaluation of comprehensive community initiatives: progress, prospects, and problems. In: Fulbright-Anderson K., Kubisch A. C., Connell J. P, eds. *New approaches to evaluating community initiatives. Vol. 2. Theory, measurement, and analysis.* Washington DC: The Aspen Institute: 15–44.

Cooper, H. M., & Tom, D. H. Y. (1984). Teacher expectation research: A review with implications for classroom instruction. *Elementary School Journal, 85*(1):77–89.

Covington, M. V. (1992). *Making the grade: A self-worth perspective on motivation and school reform.* New York: Cambridge University Press.

Firestone, W. A., & Rosenblum, S. (1988). Building commitment in urban high schools. *Educational Evaluation and Policy Analysis, 10*(4): 285–299.

Gates, B. (February 22, 2005). Speech to the National Governors' Association. Retrieved May 16, 2005 from: http://www.gatesfoundation.org/MediaCenter/Speeches/BillgSpeeches/BGSpeechNGA-050226.htm?version=print.

Greene, J. P., & Winters, M. A. (2005). *Public high school graduation and college-readiness rates: 1991–2002. Education working paper no. 8.* New York: Center for Civic Innovation, Manhattan Institute for Policy Research.

Lee, V. E., Smith, J. B., Perry, T. E., & Smylie, M. A. (1999). *Social support, academic press, and student achievement: A view from the middle grades in Chicago.* Chicago, IL: Consortium on Chicago School Research.

Marks, H. M., Secada, W. G., &Doane, K. B. (1996). *Social support for achievement: building intellectual culture in restructuring schools.* Washington, DC: Educational Resources Information Center, Office of Educational Research and Improvement, U.S. Department of Education.

McLaughlin, M. W., & Talbert, J. E. (1993). *Contexts that matter for teaching and learning: strategic opportunities for meeting the nation's educational goals.* Stanford, CA: Center for Research on the Context of Secondary School Teaching, Stanford University.

National Association of Secondary School Principals. (1996). *Breaking ranks: Changing an American institution.* Reston, VA: Author.

Newmann, F. M., Marks, H. M., & Gamoran, A. (1996). Authentic pedagogy and student performance. *American Journal of Education, 104*(4): 280–312.

Raudenbush, S. W. (1984). Magnitude of teacher expectancy effects on pupil IQ as a function of the credibility of expectancy induction. *Journal of Educational Psychology, 76*(1): 85–97.

Sizer, T. R. (1992). *Horace's school: Redesigning the American high school.* Boston, MA: Houghton Mifflin.

Smerdon, B. A. (1999). How perceptions of school membership influence high school students' academic development: Implications for adolescents at risk of educational failure. Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

Vander Ark, T. (2001). Rethinking our educational system from the classroom to the boardroom. *Business Perspectives,* Spring: 18–25.

Endnotes

[1] AIR/SRI are conducting additional achievement analyses in other districts. We included districts for which we had completed preliminary analyses.

[2] In the four districts examined here, only one school was in the planning stage of the redesign process. This "planning year" school was not included in the analysis.

[3] We did not include the performance-based promotion attribute in creating the implementation index or in our analyses, because the measure we have for this particular attribute was highly unreliable and poorly correlated with the other six attributes or the overall implementation index.

[4] For the purpose of this report, the minority group was defined as black, Hispanic, and American Indian students.

Appendix A

Methods of Quantitative Analyses for Research Question 1

Data and Sample

To address Research Question 1, we relied on extant data culled from district and/or state web sites as well as data provided by the individual districts. The number of schools included in our full sample was presented in Table1 in the text. Because data availability varies across schools and districts, we were unable to collect data for all measures for all schools. The actual number of schools used in our analyses is provided in Table A1 in the next page. The number of students in the analytic samples for achievement measures is also provided in Table A1.

Measures

Extant school-level data were used to measure school demographic characteristics and intermediate student outcomes. The demographic characteristics we examined included school minority composition and the percentage of students who qualify for free and reduced-priced lunch, special education services, and English language learner services. School-level, non-achievement outcomes included average daily attendance, suspension rates, and progression rates. Data on all these measures except progression rates were from district websites, and data on progression rates were provided by individual districts upon our request. In addition, individual districts also provided both middle school and high school student-level assessment data in ELA/Reading and math along with student ID. In some cases, assessment data in other subject areas were provided as well.

Analytic Methods

Demographic and non-assessment intermediate outcomes were analyzed at the school level, while analyses of assessment data were conducted at the student level. Our analyses of demographic and non-assessment data at the school level had only limited statistical power due to the small number of high schools in each district. Conversely, the large number of students in our analyses of achievement data at the student level provided strong statistical power and made statistical significance relatively easy to achieve. Because of the sensitivity of statistical power to sample size, we focus on the magnitude rather than the statistical significance of our results, and report those results in terms of effect sizes. We computed

Table A1. Number of Schools in Analytic Samples by District, School type, and Measure*

	District One		*District Two*			*District Three*			*District Four*	
Measures	*Redesign*	*Other*	*New*	*Redesign*	*Other*	*New*	*Redesign*	*Other*	*New*	*Other*
Percent Minority	7	17	2	11	8	4	7	7	3	6
Percent Free or reduced-price lunch	6	17	2	11	8	4	7	5	3	6
Percent Individual Education Program	7	13	2	11	8	4	7	7	3	6
Percent English Language Learners	4	12	2	10	8	4	7	7	3	6
Middle school math achievement	7 (390)	17 (2961)	2 (98)	11 (726)	8 (997)	4 (108)	7 (717)	7 (1245)	3 (173)	6 (928)
Middle school reading achievement	7 (400)	17 (3004)	2 (96)	11 (720)	8 (997)	4 (78)	7 (527)	7 (981)	3 (160)	6 (864)
High school math achievement	Not analyzed	Not analyzed	2 (97)	11 (726)	8 (990)	2 (111)	7 (683)	7 (1261)	3 (291)	6 (1474)
High school reading achievement	7 (375)	17 (2668)	2 (95)	11 (720)	8 (995)	4 (110)	7 (674)	7 (1248)	3 (291)	6 (1474)
Attendance rate	7	17	2	10	8	4	7	7	3	6
Progression rate	5	16	2	8	8	4	7	7	3	5
Suspension rate	6	16	2	11	8	4	7	7	3	6

* Note. Number of students for analyses of achievement measures is provided in parentheses.

the effect sizes of different measures by dividing the differences in means between the two types of schools compared by the pooled standard deviations derived as follows:

$$\text{Standard Deviation}_{\text{pooled}} = \sqrt{\frac{((n_1 - 1) * SD_1^{\,2}) + ((n_2 - 1) * SD_2^{\,2})}{n_1 + n_2 - 2}}$$

For school-level analyses, n1 and n2, in the above formula are the number of schools, and SD1 and SD2 are the school-level standard deviations of the two types of schools compared respectively. For student-level analyses, n1 and n2 are the number of students, and SD1 and SD2 are the student-level standard deviations of the two types of schools compared respectively.

Table 3 in the text presents the effect sizes in high school achievement as well as the effect of controlling for prior achievement. Given the nature of our research questions and the achievement data available, we employed student-level regression analyses to assess the differences between school types in students' performance on the math and reading proficiency exams in high school, while controlling for prior achievement and when available, student gender, minority status, eligibility for free and/or reduced-price lunch. We performed a separate analysis for each district as well as subject. The regression model for each district is specified below.

District One:

Y = B0 + B1*(GENDER) + B2*(MINORITY) + B3*(LUNCH) +
B4* (PRIOR ACHIEVMENT) + B5*(REDESIGN) + R

District Two:

Y = B0 + B1*(GENDER) + B2*(MINORITY) + B3*
(PRIOR ACHIEVMENT) + B4*(NEW) + B5*(REDESIGN) + R

District Three:

Y = B0 + B1*(GENDER) + B2*(MINORITY) + B3*
(PRIOR ACHIEVMENT) + B4*(NEW) + B5*(REDESIGN) + R

District Four:

Y = B0 + B1*(GENDER) + B2*(MINORITY) + B3*
(PRIOR ACHIEVMENT) + B4*(NEW) + R

Where,

GENDER: 1=female, 0=male;
MINORITY: 1=African American, Hispanic, or American Indian; 0=White or Asian/Pacific Islanders;
LUNCH: 1=eligible for free or reduced-price lunch; 0=otherwise;
PRIOR ACHIEVMENT: student's middle school achievement;
NEW: 1=new school; 0=otherwise
REDESIGN: 1=redesign school; 0=otherwise;

In the above models, the dummy variables, NEW and REDESIGN, represent new and redesign schools respectively, with other (non-foundation-supported) schools being the reference group. Further, gender, minority status, eligibility for free or reduced-price lunch, and prior achievement were centered around their respective grand means, such that the intercept B0 represents the predicted high school test score for a student with average prior achievement who attended a school with average student compositions.

Appendix B

Methods of Quantitative Analyses for Research Question 2

Data and Sample

For Research Question 2, we examined the relationships between school organizational characteristics, teaching practices, and two types of student outcomes: student attitudes and student 10th-grade achievement in reading and math. Data on school organizational characteristics, teaching practices, and student attitudes came from surveys administered to teachers, students, and school leaders in 2002, 2003, and 2004 in a sample of 44 high schools that had adequate response rates and no missing data on key variables.[1] The 44 schools included 5 model schools, 24 first-year new schools and 15 large high schools (including both preredesign and comparison schools) pooled from all three years of survey data collection (see Table B1).

Table B1. Number of Schools for Analyses of Student Attitude Outcomes by School Type and Data Collection Year

School Type		*2002*	*2003*	*2004*	*N of Pooled Sample*
Model		5	0	0	5
New		6	11	7	24
Large	***Preredesign***	8	2	2	15
	Comparison	0	0	3	

For analyses involving student achievement outcomes, we gathered school-level achievement data only for schools for which we had survey data, and combined the achievement data with survey data for those schools. The combined data set included 20 schools, of which 8 were news, 9 were preredesigns, and 3 were redesign schools.

Measures

Implementation Index. We created an implementation index to indicate the overall level of implementation of effective-schools attributes. Specifically, we first employed factor analyses to construct a set of scales that were mapped onto the six key attributes for effective high schools identified by the foundation: Common Focus, High Expectations, Personalization, Collaboration, Respect and Responsibility, and Technology as a Tool.[2]

The relevant survey items comprising each of the scales as well as the reliabilities of the scales in each survey year are listed in Table B2. Based on these scales, we created measures of the six effective-schools attributes and the overall implementation index as follows:

- Aggregate the teacher scales and student scales comprising the six attributes to the school level.
- Standardize the aggregated teacher and student scales.
- For each of the six attributes, create an attribute measure as the mean of the standardized teacher and student scales comprising the attribute.
- Standardize the six attribute measures.
- Create the implementation index as the mean of the six standardized attribute measure.
- Standardize the implementation index.

All standardized measures had a mean of zero and a standard deviation of one, with higher values indicating stronger presence of the effective-schools attributes in the school.

Student-Centered Instruction and Teacher-Directed Instruction. Based on teacher surveys, we constructed two teacher-level composite measures of instructional practices: student-centered instruction and teacher-directed instruction, and then aggregated them to the school level. The measure of student-centered instruction was based on the following four scales: active inquiry, in-depth learning, performance assessment, and technology as a tool; the measure of teacher-directed instruction was based on two scales: traditional instruction and preparation for standardized tests. The relevant survey items comprising each of the scales as well as the reliabilities of the scales in each survey year are listed in Table B3.

Student Attitudes. Based on student surveys, we created the following five measures of student attitudes:

- *Engagement-interest* — student reports of how often they asked questions or contributed in class, met with teachers, and talked about schoolwork outside of class
- *Engagement-persistence* — student reports of how often they gave extra effort on challenging assignments, got help with difficult homework, and resisted giving up when work was hard or not interesting
- *Academic self-concept* — the degree to which students felt they were good at reading, writing, learning mathematics, getting help, and working with others

Table B2. Student Survey Items Comprising the Scales Used to Measure Effective-School Attributes and Alpha Coefficients for Each Scale, by Year and Respondents

Scale	Survey Items	Reliability (α) 2002	2003	2004
	Common Focus			
Common focus	How much do you agree: · Teachers have different visions for student learning · Teachers share beliefs about what the central mission of the school should be · Teachers are committed to developing strong relationships with students · Teachers are committed to developing partnerships with parent(s)/guardian(s) for student learning · Parents and community members share vision for student learning	Teacher (.77)	Teacher (.81)	Teacher (.83)
Instructional coherence	How much agree: · Support programs are linked to curriculum, instruction, and assessments · Professional development supports the implementation of a set of common curricula, instructional strategies, and assessments · Curricula are coordinated to avoid repeating subject matter with students as they move from grade to grade · Familiar with curricula and instructional strategies used by colleagues who are also teaching my students in subject areas other than my own · Teachers have adequate opportunity to meet with one another	Teacher (.67)	Teacher (.81)	Teacher (.75)
	High Expectations			
High expectations	How much agree: Most teachers · Set high standards for teaching · Set high standards for students' learning · Make expectations for meeting instructional goals clear to students · Carefully track students' academic progress	Teacher (.89)	Teacher (.90)	Teacher (.88)
High expectations	How much agree: Teachers at school · Believe all students can do well · Have given up on some students · Care only about smart students · Expect very little from students · Work hard to make sure all students are learning	Student (.71)	Student (.76)	Student (.73)

Table B2. Survey Items Comprising the Scales Used to Measure Effective-School Attributes and Alpha Coefficients for Each Scale, by Year and Respondents—Continued

Scale	Survey Items	Reliability (α) 2002	2003	2004
	Personalization			
Personalization-social	% of students for whom you know: · Their first and last names · Their academic aspirations · Their academic background prior to this year · Their home life · Names of person/people with whom they live · Who their friends are · Their cultural and linguistic backgrounds	Teacher (.93)	Teacher (.93)	Teacher (.93)
Personalization-academic	Extent to which you help students with academic difficulties by: · Diagnosing problems the students are having · Determining how to match school resources to student needs · Gathering info to help understand students' difficulties · Helping students learn how to overcome their difficulties in ways that compensate for different learning disabilities	Teacher (.88)	Teacher (.89)	Teacher (.88)
Personalization-school action	Extent to which your school provides following help to students with academic difficulties: · Extra attention from you · Extra help from other staff member during regular school day, week, or year · Extra help from school staff outside regular school day, week, or year · Parent-teacher meetings to discuss what the school and the student's parent(s)/guardian(s) can do to help · Referrals to community organizations for assistance · Extra help from other students	Teacher (.82)	Teacher (.85)	Teacher (.81)
Personalization	How many adults in your school: · Willing to give extra help with your schoolwork if needed · Willing to help you with a personal problem · Really care about how well you are doing in school · Have helped you think about whether you are meeting the requirements for graduation · Have helped you think about what you need to do to prepare for college or a career	Student (.84)	Student (.86)	Student (.84)

Table B2. Survey Items Comprising the Scales Used to Measure Effective-School Attributes and Alpha Coefficients for Each Scale, by Year and Respondents—Continued

Scale	Survey Items	Reliability (α) 2002	2003	2004
	Collaboration			
Time to collaborate	How often have you engaged in: · Observing other teachers while they teach · Being observed by other teachers while you teach · Receiving feedback from other teachers based on their observations of your teaching · Providing feedback to other teachers based on your observations of their teaching · Coaching or mentoring other teachers or staff in your school · Co-teaching with other teachers · Diagnosing individual students' learning with other teachers	Teacher (.80)	Teacher (.81)	Teacher (.81)
Reflective professional dialogue	How often have you met with other teachers to discuss: · The goals of this school · The structure of the school day · Development of new curricula or modification of existing curricula · Teaching practices or instructional issues · General classroom administration and management	Teacher (.87)	Teacher (.86)	Teacher (.87)
Parent involvement	How often have you: · Involved parents/guardians in setting up particular learning objectives for student · Involved parents/guardians in judging student work · Provided parents/guardians with exemplars of excellent student work to demonstrate standards for good performance · Involved parents/guardians as mentors for individual students or groups of students	Teacher (.79)	Teacher (.85)	Teacher (.81)
Community resources	How often have you: · Consulted community members to better understand your students How often in your selected instructional period: · Had a guest speaker from the school's community · Discussed different cultures in your community · Took students to visit places or organizations in the community	Teacher (.59)	Teacher (.53)	Teacher (.62)

Table B2. Survey Items Comprising the Scales Used to Measure Effective-School Attributes and Alpha Coefficients for Each Scale, by Year and Respondents—Continued

Scale	Survey Items	Reliability (α) 2002	2003	2004
	Respect and Responsibility			
Respect and responsibility	How much do you agree: · Teachers feel good about parents'/guardians' support of their work · Students treat one another with respect · Relationship between students and teachers is based on mutual trust and respect · Students get teased if they take academics seriously · Student success/failure is due to factors beyond teachers' control · I can usually get through to even the most difficult students · It is the responsibility of teachers to keep students from dropping out · Teaching makes a difference in students' lives	Teacher (.73)	Teacher (.76)	Teacher (.74)
Respect and responsibility	How much agree: · Many students in this school don't respect one another · There are groups of students in this school who don't get along How many students: · Feel it's OK to make racist or sexist remarks · Feel it's OK to cheat · Feel it's OK to get into physical fights · Feel it's OK to steal things from other students · Feel it's OK to destroy or steal school property	Student (.84)	Student (.86)	Student (.82)
Collegiality	How much agree: · Teachers really don't support each other or work together · Teachers at this school trust and respect one another · Teachers, administrators, and other staff at this school model responsible behavior for students to see	Teacher (.79)	Teacher (.76)	Teacher (.76)
School climate safe	How often have you felt unsafe: · In your classes · In hallways, stairs, and bathrooms · Immediately outside the school	Teacher (.87)	Teacher (.94)	Teacher (.90)
School climate safe	How often have you felt unsafe: · In your classes · In hallways, stairs, and bathrooms · Immediately outside the school	Student (.88)	Student (.85)	Student (.86)
Orderly climate	How often has this occurred in school: · Fighting · Destroying property · Verbal bullying · Physical bullying · Cheating · Theft	Student (.91)	Student (.93)	Student (.92)

Table B2. Survey Items Comprising the Scales Used to Measure Effective-School Attributes and Alpha Coefficients for Each Scale, by Year and Respondent—Continued

Scale	Survey Items	Reliability (α) 2002	2003	2004
	Technology as a Tool			
Technology as a tool	How often do your students use technology for: · Expressing themselves in writing · Communicating electronically about academic subjects · Exploring ideas and information · Analyzing information · Presenting information to an audience · Improving computer skills	Teacher (.90)	Teacher (.90)	Teacher (.90)

- *Satisfaction-academic progress*—how well students felt they had been taught to read, write, analyze math problems, and learn on their own
- *Satisfaction-social responsibility*—how well students felt they had been taught to be responsible members of the community, respect diverse opinions, and think critically

Higher values on each of these indices indicate more positive student attitudes. The specific survey items comprising these indices are listed in Table B4.

Student Achievement. We had two measures of student high school achievement: one for reading/ELA and one for mathematics, from the 2003–04 school year. Since the original achievement measures varied across schools, we aggregated individual student scores to the school level, and then converted the school-level measures into binary measures (1 = above district mean, 0 = at/below district mean), so that the achievement measures were compatible across all schools.

Control Measures for Statistical Analyses. In our analyses on the relationships between school organizational attributes, teaching practices, and student attitudes (see the section to follow), we controlled for a number of student and school characteristics. The student-level control measures include the following:

- *Grade (GRADE):* Grade ranged from one to seven and represented the grade the student was in when surveyed.
- *Gender (SFEMALE):* Gender is a dummy variable, with female coded as 1 and male coded as 0.
- *English language exposure (NONENG):* English language exposure is represented by a dummy variable comparing students who

Table B3. Teacher Survey Items Comprising the Scales Used to Measure Teaching Practice

Scale	Survey Items	Reliability (α) 2002	2003	2004
	Student-Centered Instruction			
Active inquiry	Emphasis in instructional period: · Explore interesting topics · Guide student research/analysis In instructional period do you: · Monitor student-led discussions · Collect/organize/analyze-info/data · Decide to present what learned · Evaluate/defend ideas and views	.78	.71	.79
In-depth learning	Emphasis in instructional period: · Relating content to real life · Help students explore topics In instructional period do students: · Solve real-world problems · Research topics enough to be expert · Work on multidisciplinary projects · Participate in community projects	.70	.72	.75
Performance assessment	How often in instructional period: · Open-ended problems · Portfolios of student work · Group projects · Individual projects · Student peer reviews · Hands-on demos/ exhibitions/ present · Performance assessments	.80	.78	.78
Technology as a tool	Teaching in instructional period: · Express themselves in writing · Communicate electronically · Explore ideas and information · Analyze information · Present information to an audience · Improve computer skills	.90	.90	.90
	Teacher-Directed Instruction			
Traditional instruction	Emphasis in instructional period: · Help students with facts and procedures · Help students w/ reading/math In instructional period do you: · Lecture to class as a whole · Lead on facts/definitions/computation · Practice computations/procedures · Memorize	.72	.71	.71
Preparation for standardized tests	Emphasis in instructional period: · Cover materials in state/district standard · Cover materials on state /district tests · Testing students · Multiple-choice tests How often in instructional period: · Short-answer tests · Prepare for state /district tests · Use technology to prepare to take standardized tests	.77	.84	.80

Table B4. Student Survey Items Comprising the Scales Used to Measure Student Attitudes

Scale	Survey Items	Reliability (α) 2002	2003	2004
	Student-Centered Instruction			
Student Engagement-Interest	This school year, · I have talked to my family about what I am working on in school. · I have asked my friends for advice about something I am working on in school. · I have asked questions in class or contributed to class discussions. · I have worked with classmates outside of class or school on schoolwork. · I have asked my teachers to meet with me to talk about grades, assignments, or my work on projects.mcq	.75	.76	.75
Student Engagement-Persistence	· I got frustrated and gave up when my schoolwork became too hard. · When my schoolwork became difficult, I found a way to get help. · I gave extra effort to challenging assignments or projects. · I kept trying to do well on my schoolwork even when it wasn't interesting to me. · I tried really hard to do a good job. · I really found my schoolwork interesting. · I really did not care too much about my schoolwork.	.76	.74	.77
Academic Self-Concept	I am good at: · asking teachers for help when I get stuck on schoolwork. · working in a group with other students. · taking part in class or group discussions. · understanding what I read. · writing papers or stories. · learning math.	.69	.70	.68
Satisfaction-Academic Progress	How well has your school taught you to: · be a good reader? · speak clearly and effectively? · write clearly and effectively? · analyze and solve math problems? · learn effectively on your own with little help from others?	.84	.79	.79
Satisfaction-Social Responsibility	How well has your school taught you to: · be a responsible member of your community? · understand the rights and responsibilities of people living in the United States? · respect the opinions of people from difference backgrounds? · prepare from the work world or attending college? · think critically about ideas, problems, and current events?	.86	.83	.81

reported that a language other than English was spoken at home for any amount of time with students for whom English was the only language spoken at home.

- *Mother's educational attainment (MOMHS & MOMCOLLG):* Mother's educational attainment was measured by two dummy variables: MOMHS (1=mother did not attend school beyond high school; 0=otherwise), and MOMCOLLG (1=mother had at least some college education; 0=otherwise). The referenced group is students who "do not know" their mothers' educational attainment.
- *Race/Ethnicity (MULTIOTH, ASIAN, HISPAN, and BLACK):* Students were assigned into five mutually exclusive racial categories represented by four dummy variables. Students who indicated that their heritage was of more than one race were assigned to the MULTIOTH category. Student who reported that they were Asian, Native Hawaiian, or Pacific Islanders were grouped to the ASIAN category. WHITE was the reference category.

In addition to the above student-level control variables, we also incorporated into our statistical model the following school-level control variables:

- *Year of survey administration (Y2003 and Y2004):* The year students took the survey was represented by three categories and two dummy variables. The 2002 administration year served as the reference year, which is contrasted with 2003 and 2004.
- *School risk index (ZRISK):* The school risk index was a composite measure based on the following school demographic characteristics: percentage of students eligible for free or reduced-priced lunch, and percentage of minority students (African Americans, Hispanics, and Native Americans). The risk index was standardized such that it had a mean of zero and a standard deviation of one. Higher values of the risk index were associated with more risk-related student characteristics in schools.

Analytic Methods

We conducted two types of analyses to address Research Question 2. First, we investigated the relationships between school organizational characteristics (i.e., the implementation index), teaching practices (i.e., student-centered instruction and teacher-directed instruction) and student attitudes in a sample of 44 schools that we surveyed using the hierarchical linear modeling (HLM) method. Second, we performed *t*-tests to examine the relationships between the above survey-based measures and students' 10th-grade achievement in a sample of 20 schools that had both survey data and achievement data.

HLM analyses of the relationships between school organizational characteristics, teaching practices, and student attitudes

Given the nested nature of our survey data (i.e., students are nested within schools), we assessed the relationships between school organizational characteristics, teaching practices, and student attitudes by employing HLM, which is a statistical technique specifically designed to analyze data of a nested nature. Because schools were likely to differ systematically in student background characteristics and school demographics, which might also affect student attitudes, we incorporated into our statistical model control variables for those potential differences. Moreover, since we pooled data from multiple survey years, we also controlled for the effects that the time of survey administration might have on students' responses. Specifically, we formulated a two-level HLM model, with level-1 being student level and level-2 school level. Because in our survey data set, students are only linked to schools but not to teachers, we aggregated the teacher-level measures of student-centered instruction and teacher-directed instruction to the school level, and treated them as school characteristics in our HLM model, which is specified as follows:

Level-1 Model (Students)

$$Y_{ij} = \beta_{oj} + \beta_{1j}(GRADE)_{ij} + \beta_{2j}(SFEMALE)_{ij} + \beta_{3j}(NONENG)_{ij} + \beta_{4j}(MOMHS)_{ij} + \beta_{5j}(MOMCOLLG)_{ij} + \beta_{6j}(MULTIOTH)_{ij} + \beta_{7j}(ASIAN)_{ij} + \beta_{8j}(HISPAN)_{ij} + \beta_{9j}(BLACK)_{ij} + r_{ij}$$

Where,

- Y_{ij}: is the value of a student attitude measure (i.e., engagement-interest, engagement-persistence, academic self-concept, satisfaction-academic progress, or satisfaction-social responsibility) for student i in school j.
- β_{oj} is the average level of student attitude adjusted for student characteristics included in the model in school j.j
- β_{oj} (g = 1 ~ 9) is the effect of a student characteristic on student attitude for student i in school j, controlling for other student characteristics included in the model. The student characteristic variables were grand-mean centered and fixed at level 2.
- r_{ij}: is measurement error associated with student i in school j on the student attitude adjusted for student characteristics included in the model.

Level -2 Model (Schools)

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(Y2003)_j + \gamma_{02}(Y2004)_j + \gamma_{03}(ZRISK)_j + \gamma_{04}(SCH\_CHARACTERISTIC)_j + u_{0j}$$

$$\beta_{gj} = \gamma_{g0}, \text{ where } g = 1, \ldots, 9$$

Where,

- γ_{00}: is the average level of student attitude across all schools adjusted for both student characteristics and school characteristics included in the model.
- γ_{01} and γ_{02} are the effects of the year of survey administration (2003 vs. 2002, and 2004 vs. 2002) on school average student attitude controlling for both student characteristics and other school characteristics included in the model.
- γ_{03} is the effect of the school risk index on school average student attitude controlling for both student characteristics and other school characteristics included in the model.
- γ_{04} is the effect of the school characteristic of interest (i.e., the implementation index, student-centered instruction, or teacher-directed instruction) on school average student attitude controlling for both student characteristics and other school characteristics included in the model.
- u_{0j}: is measurement error associated with school j on school average student attitude controlling for both student characteristics and school characteristics included in the model.

We conducted three sets of analyses—one set for each of the three school characteristics (i.e., the implementation index, student-centered instruction, and teacher-directed instruction), with each set including five separate analyses—one for each of the five student attitude measures (i.e., engagement-interest, engagement-persistence, academic self-concept, satisfaction-academic progress, and satisfaction-social responsibility).

T-tests of the relationships between survey-based measures and students' 10th-grade achievement

Since the sample of schools in our study used different types of 10th-grade tests, the achievement data we obtained were not comparable across schools. To overcome the problem, we converted the original continuous achievement measure to a binary measure based on whether the average score of a school was above the district mean or below district mean. We then performed *t*-tests to assess whether schools achieving

above district mean differed significantly from schools achieving at/below district mean on measures of implementation, student-centered instruction, teacher-directed instruction, and student attitudes. Separate analyses were performed for reading and math.

Appendix B Endnotes

[1] Schools with adequate rates are schools in which both teacher and student response rates were at or above 50%, with at least one of the responses rates at or above 60%, for all types of schools except comparison schools. The corresponding criteria for comparison schools are 40% and 50% respectively.

[2] We did not include the performance-based promotion attribute in creating the implementation index or in our analyses, because the measure we have for this particular attribute was highly unreliable and poorly correlated with the other six attributes or the overall implementation index.

Appendix C

Methods for Site Visit Data Collection and Analyses

Site Visits

In 2003–04, site visits were conducted in 30 of the schools in the survey sample. Twenty-two new, 4 preredesign, and 4 postredesign school campuses were visited. Of the new schools, 8 schools that opened in fall 2001 were visited for the third time, 7 schools that opened in fall 2002 were visited for the second time, and 7 newly opened schools were visited for the first time. Initial visits were made to the preredesign schools, and postredesign schools were visited for the third time.

Two-person teams visited each school over a period of 2 to 4 days, as needed. School site visits included interviews with school principals and other leaders considered key to the success of reform activities, focus groups with two groups of students and (in selected schools) with two groups of parents, interviews with five teachers, and observations of five classrooms (where possible, those of teachers who were interviewed). Some school data collection instruments (e.g., interview protocols) were tailored to the circumstances of new, preredesign planning, and redesign sites. Interviews and focus groups were audiotaped to support the completeness and accuracy of the data records.

Principals and Lead Staff

Site visit teams began and ended school visits with principal interviews (where possible). Site visitors also interviewed reform facilitators, coaches, design team leaders, curriculum leaders, and others considered key to the success of the reform. These interviews covered topics such as conception of the school's mission, supports attributed to the grantee organization, school governance, and academic organization.

Teachers

Site visit teams interviewed five teachers at each school. Teachers to be interviewed were selected to meet the following criteria at schools where the criteria were consonant with the structure of the staff:

- A 10th-grade mathematics teacher (if the school didn't have a 10th-grade mathematics teacher, we interviewed a 9th- or 11th-grade mathematics teacher).
- A 10th-grade English/language arts teacher (if the school didn't have a 10th-grade English/language arts teacher, we interviewed a 9th- or 11th-grade mathematics teacher).
- A teacher of any subject at the 9th-grade level (if the school didn't have a 9th grade, we selected a teacher in the lowest grade above 8th grade).
- A teacher of any subject at the 11th- or 12th-grade level (if the school didn't have an 11th or 12th grade, we selected a teacher at the school's highest grade).
- Someone who taught an innovative class (e.g., service-learning, career course, student advisory, etc.), preferably at a higher grade level in the school.

These categories were incongruent with the school structures of some of the schools, particularly the model schools. For example, some schools do not have discrete English/language arts or mathematics classes; in these cases, we asked leaders to identify teachers of classes where mathematics and English/language arts were substantial parts of instruction. Some of these schools do not group students by grade level in mathematics and language arts. In these cases, we selected teachers so that their five classes represented a range of student levels. Site visit teams tried to schedule teacher interviews so that the same teachers could be part of the classroom observations (see below). In addition to topics addressed by school leader protocols, teacher interviews probed for relationships among teachers and between teachers and students, the school's learning environment, and the school's ability to serve all students well.

Students

Site visit teams completed two student focus groups per school. Students were taken from the classes of teacher interviewees, when possible, with one six-member group coming from one of the lowest-grade classes in the school and one from one of the highest-grade classes. Schools were asked to select from among the more heterogeneous of these classes. Selected classes were asked to take parent consent forms home for parent signature, and focus group students were selected from among those who returned signed forms. School coordinators were asked to select a mix of students by gender, racial/ethnic group, and native language status for each group. In focus groups, students were asked to describe how their school is different from or similar to other schools, the nature

of relationships among students and between students and teachers at the school, the nature of their schoolwork, and their assessment of how well the school is preparing them for life after graduation.

Classrooms

The site visit teams conducted 25-minute observations in the classrooms of interviewed teachers. Structured observation forms were used to code the structure of the instructional activity, teacher actions, and student actions. The instructional activity codes indicated how the teacher and students were grouped for teaching and learning—for example, whether the teacher was lecturing to the whole class, students were working individually, or students were working in small groups. Teacher action codes captured the role of the teacher within the activity, that is, whether the teacher was giving directions, posing questions, leading discussions, monitoring student work, and so on. Student action codes indicated what the observed students were doing, that is, whether students were listening, reading, collecting data, writing, performing, and so on. In another section of the observation form, observers provided a narrative description of the activities they had observed. The form also required observers to note the instructional resources used and aspects of classroom management, such as the proportion of students who were "on task" during the activity. After the observations, visitors met with teachers to discuss what they had seen. Observers asked teachers whether the work they had observed was part of a long-term product and, if so, whether students were using rubrics to examine their work, whether students would have opportunities to revise their work, and whether students would have opportunities to apply what they had learned to real-world contexts.

Buildings and Structures

At the conclusion of the school visit, site visitors completed an Implementation and School Environment Inventory. The inventory described the physical environment of the school, catalogued the school design components that were in planning or in place, noted the correspondence between the school model and school environment, and described the school location and neighborhood.

Data Coding

After returning from visits to schools, site visitors organized the data they had collected into data capture forms. For each type of interview, there was a form with a set of headings, organizing the data in a structure

parallel to the flow of the interview protocol. In addition, a school summary form was used to capture more general or synthetic impressions. Site visitors completed the data capture forms on the basis of their notes, checking interview tapes when appropriate for clarification or to obtain exact wording for quotations. Conventions were used to indicate the source of each piece of information, to designate the speaker's exact words as opposed to paraphrases, and to distinguish between data that came directly from the interview and inferences or clarifications provided by the site visitor. Senior analysts reviewed the data capture forms and requested clarifications and additions as needed.

In preparation for data coding, we developed a manual of codes, definitions, and procedures. Codes were developed for the constructs in the foundation's theory of change and for additional constructs in the conceptual framework. Codes described capacity issues, key school attributes, characteristics of curriculum and instruction, learning outcomes, other student and school outcomes, and many other topic areas. Each of these broad coding categories included codes for subtopics. Codes were designed to allow parsing of data capture forms by topic, so that data on similar topics across interviews could be analyzed as a set. There were 132 codes in all.

Data coding began with test coding, moved on to reliability and validity coding, and concluded with operational coding. After the coding structure used with 2003 data was refined, nine coders were trained to use the new draft coding manual and worked in pairs on a sample set of data capture forms to test the codes. Throughout the test coding process, weekly meetings among the coders and several analysts offered an opportunity for joint review of coding results and discussion of potentially ambiguous codes or other needed revisions to the coding manual.

Once the coding structure was tested and refined, subsets of five or six data capture forms at a time were selected to cover a wide variety of form types and content areas. These data forms were used to conduct reliability and validity trials. The trials were designed to promote common uses of codes across coders and to ensure that segments of text were coded as analysts would expect. Coders coded the text segments individually. The submitted individual coding choices were reviewed by two senior analysts, who then developed a set of master codes for the main ideas of the paragraph that were negotiated with the coding team. The resulting set of codes, agreed on by coders and analysts, was taken as the standard against which coders' original individual responses were compared to examine the reliability and validity of coding decisions. Agreements and disagreements with the standard codes for each paragraph's main ideas were tallied by code, and agreement was calculated as agreements/

(agreements + disagreements). In a meeting, the reasons for any low scores were explored and other outstanding issues were resolved. The coding definitions were then updated to improve clarity where necessary, and the process was repeated with the new set of definitions.

An initial reliability run was conducted to verify that each coder was sufficiently trained for operational coding to begin. During operational coding, the reliability process was repeated several times at 2-week intervals to develop our final sample for reliability and validity. In the cumulative sample from three reliability runs, 80% of codes that were used more than three times in the coding sample had estimated reliabilities ranging from 75% to 100%. Codes below that threshold generally corresponded to concepts that were difficult to separate from related topics in the narratives. For example, issues of common focus among teaching staff were often discussed in the same breath as schoolwide professional development sessions that often included discussions of the school mission and goals. Interrelated constructs like this made coding distinctions challenging. In cases like these, we computed reliability estimates for two interrelated codes together and encouraged analysts to consider using queries of both codes when they conducted analyses on these topics.

Once we moved from reliability to operational coding, weekly meetings continued for the resolution of any new issues that arose. To the extent that these discussions resulted in changes to accepted coding definitions, coders were asked to go back to previously coded documents to implement the changes.

Data Analysis

Many of the analysts of school-level data began their work by reviewing samples of data capture forms for schools in their analysis group. These reviews helped analysts get a more comprehensive view of the school contexts and schoolwide issues.

School-level analysts then queried the ATLAS.ti database to review coded data by topic. In some cases, they used coded data to find examples of issues that surfaced in survey analysis. More often, however, they used the narrative data to surface and substantiate the most prevalent themes in the coded data and to confirm or disconfirm findings suggested by the survey data. To accomplish these aims, analysts consulted the coded data on each topic, generated an initial set of themes to pursue, and developed matrices and other supporting documents to track whether or not, and in what way, a particular issue was in evidence at each school. To vet and refine the emerging themes, analysts worked in small teams by topic area and iteratively reviewed and discussed data until they

reached consensus on the supported themes. A larger team of qualitative and quantitative analysts met weekly to evaluate the qualitative themes and examine the consistency of findings across the qualitative and survey data and to decide on areas that warranted further analysis.

Appendix D

Table D1. Effect of Implementation on Student Attitudes

Student Attitude	*Coefficient*	*Standard Error*	*T- statistic*	*P-value*
Engagement-Interest	.309*	.027	11.367	.000
Engagement-Persistence	.266*	.030	8.830	.000
Academic Self-Concept	.181*	.017	10.684	.000
Satisfaction-Academic Progress	.237*	.020	11.686	.000
Satisfaction-Social Responsibility	.317*	.033	9.519	.000

Note: Number of schools = 44; Number of students = 10887. * = p<.10

Source: Data from the school information form, teacher and student surveys from 2001–02, 2002–03 and 2003–04.

Table D2. Effect of Student-centered Teaching on Student Attitudes

Student Attitude	*Coefficient*	*Standard Error*	*T- statistic*	*P-value*
Engagement-Interest	.283*	.036	7.829	.000
Engagement-Persistence	.289*	.028	10.253	.000
Academic Self-Concept	.169*	.019	9.024	.000
Satisfaction-Academic Progress	.228*	.023	9.800	.000
Satisfaction-Social Responsibility	.289*	.043	6.766	.000

Note: Number of schools = 44; Number of students = 10887. * = p<.10

Source: Data from the school information form, teacher and student surveys from 2001–02, 2002–03 and 2003–04.

Table D3. Effect of Teacher-directed Teaching on Student Attitudes

Student Attitude	*Coefficient*	*Standard Error*	*T- statistic*	*P-value*
Engagement-Interest	-.273*	.031	-8.754	.000
Engagement-Persistence	-.197*	.042	-4.709	.000
Academic Self-Concept	-.142*	.025	-5.659	.000
Satisfaction-Academic Progress	-.184*	.029	-6.452	.000
Satisfaction-Social Responsibility	-.263*	.041	-6.326	.000

Note: Number of schools = 44; Number of students = 10887. * = p<.10

Source: Data from the school information form, teacher and student surveys from 2001–02, 2002–03 and 2003–04.

Table D4. Results of T-tests Comparing Schools Achieving above District Mean and Schools Achieving at/Below District Mean on the Implementation Index, Student-Centered Instruction, and Teacher-Directed Instruction

School Characteristics		*T-statistic*	*Standard Error*	*P-value*
Implementation Index	ELA	2.462*	.383	.026
	Math	.269	.539	.791
Student-Centered Instruction	ELA	1.102	.431	.286
	Math	-.975	.530	.345
Teacher-Directed Instruction	ELA	-1.416	.467	.176
	Math	.079	.538	.938

Note: Number of schools = 19 for Reading/ELA, and 17 for math. * = p<.10

Source: Data from the school information form, teacher and student surveys from 2001–02, 2002–03 and 2003-04 and district provided assessment data for 2003–04.

Table D5. Relationship Between Student Attitudes and Student ELA Performance

Student Attitude		*T-statistic*	*Standard Error*	*P-value*
Engagement-Interest	ELA	2.926*	.342	.01
	Math	-.719	.469	.483
Engagement-Persistence	ELA	1.005	.392	.331
	Math	-.693	.536	.498
Academic Self-Concept	ELA	1.676	.397	.111
	Math	-.779	.501	.448
Satisfaction-Academic Progress	ELA	-.007	.498	.994
	Math	-.917	.431	.373
Satisfaction-Social Responsibility	ELA	.459	.451	.656
	Math	-.614	.418	.548

Note: Number of schools = 20 for Reading/ELA, and 18 for math. * = p<.10

Source: Data from the school information form, teacher and student surveys from 2001–02, 2002–03 and 2003-04 and district provided assessment data for 2003–04.


National
Evaluation of
High School
Transformation