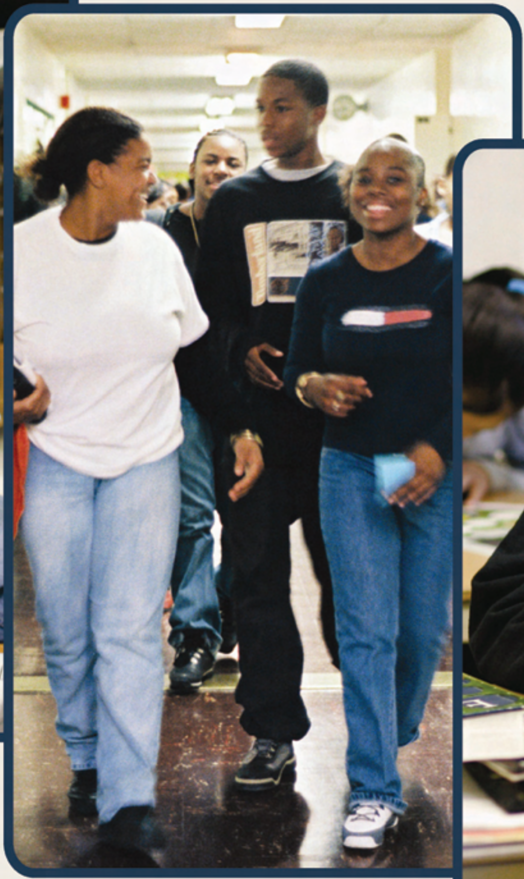
Evaluation of the Bill & Melinda Gates Foundation's High School Grants Initiative

2001-2005 Final Report



Prepared by:

The American Institutes for Research

SRI International

AMERICAN INSTITUTES FOR RESEARCH®

SRI International®

Evaluation of the Bill & Melinda Gates Foundation's High School Grants Initiative: 2001-2005 Final Report

August 2006

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 Street, NW
Washington, DC 20007
http://www.air.org

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

The work described in this document was supported by the Bill & Melinda Gates Foundation through an evaluation contract. Opinions expressed herein are the authors' and do not necessarily reflect those of the foundation. Please direct inquiries to Charles Storey at The American Institutes for Research, 1000 Thomas Jefferson Street, NW, Washington, DC 20007 or cstorey@air.org.

The Evaluation of the Bill & Melinda Gates Foundation's High School Grants Initiative 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.

This Final Report is the final installment 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

Aimee Evan
Mette Huberman
Barbara Means
Karen Mitchell
Linda Shear
Jamie Shkolnik
Becky Smerdon
Mengli Song
Charles Storey
Kazuaki Uekawa

Principal Investigators

Becky Smerdon
American Institutes for Research

Barbara Means
SRI International

Achievement Study Team

Andrea Berger
Winona Burt
Aimee Evan
Matt Hickey
Amber Noel
David Rhodes
Scott Richman
Adam Segal
Mengli Song
Charles Storey
Tyler Vogel

Rigor and Relevance Team

Jennifer Anthony
Ashley Fedor
Matt Gaertner
Larry Gallagher
Mike Garet
Amy Hafter
Lauren Hersh
David Joliffe, University of Arkansas
Joel Knudson
Steven Leinwand
Connie Liu
Carmen Manning, University of Wisconsin, Eau Claire
Karen Mitchell
Robert Murphy
Jamie Shkolnik
Mengli Song
Kazuaki Uekawa

District Portfolio Team

Winona Burt
Susan Cole
Lenay Dunn
Torie Gorges
Ann House
Mette Huberman
Aasha Joshi
Karen Mitchell
Gabrielle Phillips
Linda Shear
Pam Stoddard
Melinda Tracey
Edith Yang

Contents

List of Exhibits

List of Figures

List of Tables

Chapter 1. Introduction

The Bill & Melinda Gates Foundation has set no less a goal than redefining the American high school. Bill Gates, Jr., in a highly publicized address before the National Education Summit on High Schools, declared:

> America's high schools are obsolete. By obsolete, I don't just mean that our high schools are broken, flawed, and under-funded—though a case could be made for every one of those points. 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 . . . Today, only one third of our students graduate from high school ready for college, work, and citizenship. The other two thirds, most of them low-income and minority students, are tracked into courses that won't ever get them ready for college or prepare them for a family-wage job—no matter how well the students learn or the teachers teach. This isn't an accident or a flaw in the system; it is the system. (Bill & Melinda Gates Foundation, 2005b)

This report describes the results associated with the foundation's high school grants initiative. Under the initiative, grants were made to intermediary organizations charged with starting new high schools or redesigning existing schools in ways that better prepare low-income, African-American, and Hispanic/Latino youth for higher education and the workplace. Outcomes of this program, begun in 2000, need to be understood in the context of a broader, 10-year philanthropic and policy advocacy strategy.

In 2001, the foundation funded a 5-year evaluation project to be conducted by the American Institutes for Research (AIR) and SRI International. The research team was charged with examining the efforts of a set of organizations receiving grants to support the reinvention of high schools in locations outside the state of Washington.[1] The evaluation focused on grants involving either the creation of new high schools or the redesign of existing high schools, typically into smaller independent schools or learning communities sharing an existing campus. As the foundation awarded additional grants under this initiative, it directed the evaluation team to incorporate the most critical grants into the data collection and analysis activities. Eventually, the AIR/SRI data collection encompassed 22 grantee organizations starting or redesigning hundreds of secondary schools. The major data sources for the evaluation have been student, teacher, and principal surveys; teacher and principal interviews; student and parent focus groups; classroom observations; samples of teacher assignments and associated student work; interviews with district and grantee organization staff; and, in selected districts, analyses of test scores and behavioral data for all high schools in the districts.

[1] A series of evaluation reports on the foundation's Washington State high school grants has been prepared by Fouts & Associates (available at http://www.gatesfoundation.org/education/researchandevaluation/evaluation/msdgevaluation.htm).

Qualitative and quantitative data were collected for four model schools being replicated in multiple locales. Data were also collected for samples of new schools and for large schools planning their redesign, as well as the schools or learning communities resulting from the redesign process. Data were collected in multiple years for all except the model schools. The teacher assignments and student work, along with test scores for the same classes, were collected from a subset of the schools participating in site visits and survey data collections. We also collected data from a set of "comparison schools"—the schools that students attending new schools were likely to have attended if the new schools had not been an option. Finally, we examined publicly available information and gathered extant data from a subset of districts in which the foundation has helped to create or redesign high schools.[2]

An Evolving High School Reform Strategy

From its beginning in 2000, the national high school grants program represented a balance between pragmatism and idealism. Some of the grants called for intensive work with particular districts, including Oakland, Cincinnati, and Providence. These grants typically went to external nonprofit organizations that sought to help the districts redesign failing urban schools. District involvement enabled resources to be leveraged for the effort and created the potential to affect large numbers of students, but it also sometimes entailed constraints imposed by district policies, practices, and limited appetites for change. Other early grants went to organizations that would start new "break-the-mold" schools. Some of these organizations had already started small schools that had proven successful in serving diverse student bodies; in essence, the organizations had existence proofs for the efficacy of their school models. These model schools—The Met in Providence, Minnesota New Country, New Tech High in Napa, and High Tech High in San Diego—embodied a progressive approach to secondary education. Grants from the Bill & Melinda Gates Foundation would provide the opportunity to replicate these model schools in multiple communities across the country. Other organizations receiving grants to start new high schools had never before created a secondary school but worked instead from a set of design principles or a process for supporting teachers and communities in developing school designs (AIR/SRI, 2003).

Many of the leaders of the first organizations receiving grants had close ties to Ted Sizer (the founder of the Coalition of Essential Schools) and embraced his philosophy of emphasizing a small set of essential habits of mind rather than the "mile-wide, inch-deep" curriculum typical of American high schools (AIR/SRI, 2003). Although the small size the foundation recommended for high schools (not more than 100 students in a grade) was what attracted the attention of outside observers, the foundation's education initiative was never about small size per se. Foundation staff identified a set of characteristics—or "attributes"—commonly found in high schools that successfully

[2] These districts were Anderson Union, CA; Baltimore, MD; Chicago, IL; Cincinnati, OH; Cotati-Rohnert Park, CA; Denver, CO; El Dorado, CA; Milwaukee, WI; New York City (Bronx, Brooklyn, and Manhattan), NY; Oakland, CA; Oklahoma City, OK; Providence, RI; Ravenswood, CA; Sacramento, CA; St. Paul, MN; West Clermont, OH; and Worcester, MA.

retained students from historically underserved populations and helped them attend and graduate from institutions of higher education.

The essence of the foundation's theory of change was that the creation of high schools with these attributes would lead not only to better outcomes for the students attending the schools but also to increased demand for more such schools (AIR/SRI, 2003). A small school size was deemed a necessary but not sufficient condition for creating the desired learning environment: close, personalized relationships between students and faculty and the individualized instructional program and tailored assistance needed to motivate and enable students' high performance (see Table 1).

Table 1. 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.
Personalization	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 (less 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 curricula 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 dialog about continuous improvement.

Source: Bill & Melinda Gates Foundation (n.d.-a).

During the course of the evaluation, the foundation's high school reform strategy has undergone a significant evolution. This evolution was prompted by insights acquired from the foundation's experiences with its early grants, changing external conditions, and an increase in the education program's staff capacity. Additionally, early evaluation findings highlighted both strengths and weaknesses within the initiative. Important findings include the following:

- New schools that were opened with support from organizations that received foundation grants are characterized by dramatically greater personalization, higher expectations of students, and a more cohesive teacher community than are found in large, comprehensive high schools.
- Students in the new schools have more positive educational attitudes; they feel more supported by their teachers and they are more interested in their schoolwork than students in more conventional high schools. They also have higher attendance and, although they enter high school with lower achievement levels than other students in their districts, generally make progress relative to district averages in English/language arts.
- In their second year, new schools typically experience "growing pains," with some erosion in the strength of their school climate and signs of teacher burnout, as they continue to try to provide a wide range of services and student-centered instruction in addition to adding a new grade level.
- Existing schools planning a redesign to better reflect the attributes of high-performing schools need more than a single year for the design process. After they go through the redesign, the resulting small schools or learning communities experience positive changes in the level of personalization and sense of community, but these do not rise to the level found in new schools created from scratch.
- New schools struggle with recruiting the right kind of staff and with limitations in their funding. Schools undergoing redesign struggle with issues of changing their physical and organizational structure, defining distinctive programs for the subunits resulting from redesign, and assigning students and teachers to subunits in ways that provide both equity and motivation.
- Both new and redesigned schools need more help with issues of curriculum and instruction. Mathematics proves especially challenging for new schools, which have few staff members in total and have particular difficulty recruiting qualified mathematics teachers.

As foundation officials expected, putting the attributes of a high-performing school in place in a new school was easier than grafting the attributes onto existing schools. At existing schools, entrenched cultures and sets of expectations about student achievement and behavior often became obstacles. Dramatic differences in climate between new small schools and conventional comprehensive schools were confirmed, as well, with new schools enjoying clear strengths in terms of students' engagement with academics and their school communities.

Areas where early experiences were less positive led the foundation to refine its theory of change. Assumptions concerning the nature of systems change, the degree of emphasis on continuous monitoring of outcomes, and the stance toward curriculum and model specification were all revised over time. The foundation began stressing the importance of (a) a clearly specified and controlled "tight" model of innovation coupled with extensive supports for new schools, and (b) early attention to issues of curriculum and instructional practice in schools undergoing redesign.

Impact of the No Child Left Behind Act

During the first 3 years of the initiative, there were also important developments in the national education picture. The passage of the No Child Left Behind Act of 2001 (NCLB) increased the accountability pressures on all public schools, including high schools. NCLB requires states to set adequate yearly progress (AYP) goals for their high schools, focusing specifically on student achievement and graduation rate goals. Consequences for a school's failure to meet AYP goals in these two areas are severe and include (1) giving the school's students the choice of attending other public schools, (2) replacement of school staff, and (3) forced restructuring of the failing school. Because the legislation requires schools to disaggregate their test data by student demographic group, the performance of low-income, African-American, and Hispanic/Latino students can no longer be ignored, even in schools and districts with high mean test scores. In this way, the emphasis of NCLB converged with that of the Bill & Melinda Gates Foundation education initiative.

In other ways, however, NCLB has complicated the reform activity. State and district responses to the NCLB legislation have generally gone beyond merely testing and disaggregation of student test scores. Responses have also included increased top-down emphasis on curriculum standards and requirements that schools provide instruction based on those standards. At an extreme, some districts have provided teachers with pacing charts showing where in a curriculum classes should be working each month and even each week of the school year. All of these changes have made it more difficult for high schools to take a radically different approach to curriculum and instruction.

NCLB also requires that all high school teachers of core academic subjects be highly qualified (generally interpreted as possession of full certification, a bachelor's degree, and demonstrated competence in the subject area) by the year 2005-06. Meeting these requirements can be challenging for new schools with small staffs, especially for schools that recruited their original set of teachers for their youth development skills rather than their disciplinary focus.

Evolution of the Foundation's Theory of Change

Finally, the thinking of the foundation's education program staff has evolved in ways that have influenced the foundation's strategy. After examining projections for the total number of students they could support through grants for the creation of high schools serving 100 students per grade, foundation education staff concluded that a refinement in their strategy was needed. Staff then modified their education strategy to increase the relative weight given to efforts to redesign failing existing schools. They also chose to concentrate efforts and resources in a handful of strategically chosen cities and to work more closely with policymakers to influence their actions and investments.

During this period, the foundation's education program was also greatly expanding its staff. When the national high school grants initiative was started, the education program had roughly a half-dozen staff members to plan, run, and evaluate a $350-million initiative. As the staff dramatically increased in size, it added expertise in education policy and social entrepreneurship. With a larger staff, the program was able to extend its involvement with grantees, districts, and policymakers in the places where it had the largest investments. The foundation staff also articulated a specific, measurable target—an 80% minority student high school graduation rate by 2014—and developed an orientation around continuous measurement of outcomes. The revised theory of change therefore deemphasized promotion of a particular school size, structure, or instructional philosophy in favor of an increased focus on outcomes for high-need students.

The new theory of change also emphasized developing a system-level strategy of more direct work with districts and heightened activity on the policy front. This new emphasis recognized the important role a district plays in supporting systemwide improvement in graduation rates. Thus, over time the initiative has moved from promotion of school models on the basis of their good ideas to a more results-oriented pragmatism. Rather than look for individual project champions within a school or in a superintendent's office, the foundation looks for well-explicated, "tight" models that can guide the reform even if a charismatic leader leaves the organization or is not available at a replication site. The notion, articulated in The Bridgespan Group's *Expanding the Supply of High Quality Public Schools* report, was that the greater the degree of management support and design specificity a model had, the greater the likelihood that the model could be replicated in new locations (Colby, Smith, & Shelton, 2005).

The foundation also developed the concept of a portfolio of schools, offering different curricula and instructional approaches, some run by the district and some by external providers or charter management organizations. The portfolio strategy allowed the foundation's education program to resolve the tension between progressive and traditional models of curriculum and instruction. In some cases, districts implement a model with top-down control of curriculum; in other cases, outside organizations implement models with a range of instructional approaches. The intended result is that students and families of all income levels will have the range of choices previously available only to the affluent.

When the Bill & Melinda Gates Foundation decided to focus its education initiative on the reform of secondary education, it did not expect a smooth, straight path to the desired end. Foundation CEO Patty Stonesifer, extrapolating from her experience in the high-tech industry, expressed the expectation that the foundation would try multiple approaches and would use early results to weed out those that did not appear promising. As in the technology business, some false starts are to be expected. Rather than something to be ashamed of, such efforts are viewed as a natural part of the process of identifying effective strategies.

Evaluation Findings in Context

The findings in this report should be considered in light of the amount of time required for significant education reform. The schools undergoing redesign were seeking a fundamental shift in the thinking, expectations, and practices of students and teachers. Prior research suggests that such a change in mind-set and practice takes at least 3 to 5 years to implement deeply (Berends, Bodilly, & Kirby, 2002; Borman, 2005). New schools had the advantage of being able to recruit staff and students who were attracted to a distinctive vision of a rigorous high school, but most dealt with the complexity of getting a school off the ground quickly by opening with a single grade level—typically a class of ninth graders—resulting in an extended timeline for implementing a full 4-year high school program. Helping students—many of whom entered with skills three or four grade levels below their nominal grade—become college ready is a major challenge, and one not typically surmounted in 1 or 2 years. Further, with shoestring funding based on their small enrollments and before having a graduating class that could establish a record of college admission, many of the new schools have struggled to build a reputation for success.

Readers of this report also should keep in mind the eclectic and evolving nature of the foundation's education strategy. The school samples used in data analyses reported here are a fairly even balance between schools started under the early grants and those started or redesigned later in the initiative. Analyses involving schools that have been open for 4 years, however, draw on that portion of the school sample developed under the early grants. Another important distinction to keep in mind is the difference in governance structure for the various school types. Schools planning redesign and the entities resulting from their restructuring are all district-run schools, whereas the majority (though not all) of the new schools are operated as charter schools.

In this report, the last in the series, we address the question of what difference the foundation's high school initiative has made for students, schools, and school systems.[3] We use the experiences of this remarkable set of schools to draw implications for continuing work in high school reform.

Overview of This Report

The next chapter in this report presents analyses of the outcomes obtained for students in the sample of foundation-supported high schools participating in the national evaluation. The chapter examines achievement test scores, the quality of student work, attendance, and progression through grade levels. Analyses of student attitudinal

[3] Other AIR/SRI evaluation reports are available at http://www.air.org and at http://ctl.sri.com, as well as at the following foundation Web pages: http://www.gatesfoundation.org/Education/ResearchAndEvaluation/Evaluation/HSEvaluation.htm and http://www.gatesfoundation.org/Education/ResearchAndEvaluation/Evaluation/NHSDGEvaluation.htm.

measures are presented also. Chapter 3 examines the nature of instruction provided in schools started under the high school initiative. The primary data for the chapter were developed by collecting and systematically scoring samples of 10th-grade mathematics and English/language arts assignments. Having presented the evidence concerning the difference that these schools make in terms of instructional experiences and outcomes for students, the report turns to the issue of the scalability and sustainability of these reforms. Chapter 4 describes the breadth of school choice options in districts where the foundation has made a significant investment and early signs regarding the prognosis for new school models to endure beyond their foundation grant period. Chapter 5 concludes the report by articulating some key themes derived from the data and a set of recommendations for this and other reform initiatives. Details of the methodology and analyses are presented in the technical appendix.

Chapter 2. Promoting Student Success

This chapter examines the extent to which the foundation-supported schools have made progress in helping students succeed. These schools include schools that the foundation has helped create from scratch ("new schools") and schools that were created when existing high schools were redesigned into smaller learning communities ("redesigned schools"). The chapter specifically describes students' attendance, progression rates, performance on standardized tests, engagement aspirations and motivation, and quality of schoolwork. The foundation has been working with the schools in this study for 1 to 4 years as of 2004-05. Many of the foundation-supported schools have therefore not been in existence long and have not had sufficient time to systematize their practices and mature as institutions. Thus, this chapter documents the student outcomes so far in foundation-supported schools but cannot provide a summative evaluation of the schools' ultimate impact and is not a definitive exploration of student achievement in the foundation-sponsored schools. More definitive conclusions about student progress will become possible only after the schools and students have been monitored longer.[4]

Key Findings

At this stage in the schools' development, key findings of student outcomes include the following:

- More than 80% of the foundation-supported new schools in our sample had attendance rates that were higher than their district averages, and more than 60% had 9th-to-10th-grade progression rates that were higher than their district averages. The attendance and progression rates at foundation-supported redesigned schools, however, remained below district averages.
- In the studied districts, the raw average achievement test scores reported by most new and redesigned schools were below district averages on standardized tests in English/language arts (ELA) and mathematics. Although most of both types of schools had average test scores below their respective district averages, the percentages of new schools with average achievement test scores above district averages in the two subjects were higher than the percentages of redesigned schools with above-average scores.
- In a very small sample of districts and schools, more than 50% of the new and redesigned schools made gains in ELA proficiency relative to their respective districts from academic year 2002-03 to academic year 2004-05. Gains also were made in mathematics, but the percentage of schools that made gains fluctuated widely from 1 year to the next. Thus, although the foundation-supported schools had average achievement test scores that were lower than their district averages,

[4] Chapter 5 of this report makes specific recommendations for the continued tracking of progress at these schools as the foundation's initiative moves forward and as reforms in the schools continue to mature.

over time the gains in the schools' proficiency rates appeared to be greater than gains in their districts overall.

- On the basis of student work collected for the 2003-04 and 2004-05 school years, students in foundation-supported new high schools in our sample did higher-quality work in ELA classes than students in traditional high schools without foundation ties. The quality of students' work in mathematics classes in the foundation-supported new high schools was similar to that of students in other schools.[5]
- In 2004-05, students in foundation-supported new high schools in our sample had higher aspirations for their futures than students in other district high schools. Students in the new schools reported plans to graduate from high school and apply to college at higher rates than students in other schools. They also reported greater levels of engagement (persistence and interest) in their studies.

Table 2 summarizes the comparison of student outcomes in foundation-supported new and redesigned schools with district averages.[6]

Table 2. Student Outcomes in Foundation-Supported New and Redesigned Schools, Compared With District Averages

	New Schools	Redesigned Schools
Attendance	+	–
9th-to-10th-grade progression rates	+	–
Achievement test scores	–	–

Foundation-supported new schools had higher attendance rates and 9th-to-10th-grade progression rates than similar schools within the same districts. This success was not replicated with regard to achievement test scores or by any redesigned schools. Compared with schools enrolling similar student populations in the same or nearby districts, foundation-supported new schools had attendance rates and 9th-to-10th-grade progression rates that were higher than the district averages. Foundation-supported redesigned schools, however, had attendance rates and 9th-to-10th-grade progression rates that were below district averages. Achievement test scores in both types of schools were generally below district averages.

This chapter explores these and other findings. The remainder of this chapter is divided into three sections: (1) methods, (2) discussion of key findings, and (3) implications. The methods section provides a summary of the sources of data, as well as a brief description of the methods of analysis used.[7] This is followed by an elaborated presentation of the

[5] We are still collecting student work from redesigned schools, so these schools are not included in this discussion. Data on student work from redesigned schools will be presented next year.
[6] The data supporting these findings were not adjusted to control for students' prior achievement or demographic backgrounds.
[7] A more detailed discussion of methods and analytic approaches can be found in the technical appendix at the end of the report.

key findings. The chapter concludes with a discussion of the implications of these findings for the foundation's ongoing initiative.

Methods

This chapter first looks at the demographic background of students in the foundation-supported schools and then examines a variety of student outcome measures, including average daily attendance rates, 9th-to-10th-grade progression rates, test scores, and the quality of student work. The chapter also looks at students' attitudes toward schooling, reviewing measures of their interest in school, persistence, academic self-concept, and educational aspirations. To investigate these outcomes, we have drawn on a number of different data sources, including extant district demographic and achievement data maintained by school districts; student work artifacts; surveys of students; and site visits to schools, district offices, and the foundation's grantees (see Table 3).

Table 3. Types of Outcomes

Outcome	Academic Years for Which Data Are Reported	Source	Comparison Made	Statistical Method
Student population served	2003-04	Extant data from nine districts	New vs. redesign vs. district	Descriptive statistics
Average daily attendance rates	2002-03 and 2003-04	Extant data from six districts	New vs. redesign vs. district	Descriptive statistics
Ninth-to-10th-grade progression rates	2002-03 and 2003-04	Extant data from four districts	New vs. redesign vs. district	Descriptive statistics
Student engagement—Interest	2004-05	Student surveys	New vs. redesign	Hierarchical linear modeling (HLM)
Student engagement—Persistence	2004-05	Student surveys	New vs. redesign	HLM
Student academic self-concept	2004-05	Student surveys	New vs. redesign	HLM
Student educational aspirations	2004-05	Student surveys	New vs. redesign	HLM
Achievement scores	2003-04 and 2004-05	Extant data from 11 districts	New vs. redesign vs. district	Descriptive statistics
Quality of student work	2004-05	Student work from 18 new schools and 22 traditional large schools	New vs. comparison	Score created using Many-Facet Rasch Measurement modeling; comparisons made using HLM

In addition to collecting data from several sources, we also used various samples of schools in our analyses. We started all of our analyses by focusing first on seven school districts: Baltimore, Chicago, Cincinnati, Los Angeles, New York (Bronx, Brooklyn, and Manhattan only), Oakland, and Providence. These districts, in addition to hosting several foundation-supported schools, were also districts with which we had some familiarity as a result of our studies on student achievement in the individual districts. In most instances, only the districts that we worked with were able to provide us with data for analyses for both the 2002-03 and 2003-04 school years.[8]

Although the sample sizes sometimes differed, our criteria for selecting particular samples remained constant. We always used a sample size that was consistent with the purposes of our comparisons and analyses. Sample size was also determined partly by the availability of data; working with school districts across the country, availability frequently varied from district to district. The selection of the analytic sample also took into account the potential dependence among the data. In conducting regression-based analyses, for instance, we did not include both preredesign and redesigned schools in the same analysis, to help ensure the independence of observations and therefore the reliability of the results from the analyses. Whenever possible, we increased sample sizes to increase the statistical power of our analyses.

Serving Targeted Students. Our overall goal in this chapter is to focus our within-district analyses on districts where the foundation supported several schools and from which we could get the necessary data. Student-level demographic data from individual schools were available in only nine districts. Therefore, our analyses of whether the foundation was serving the students it had hoped to target were conducted with data from those nine districts only.

Attendance. Only six districts had foundation-supported schools old enough to provide attendance rate data for both the 2002-03 and 2003-04 school years. Of these six districts, only four were able to provide reliable 9th-to-10th-grade progression rate data for these 2 years. As a result, our sample sizes for these analyses were six and four districts, respectively.

Student Attitudes. Student attitudes were measured by using a national sample of schools surveyed as part of our evaluation. The sample included 40 schools: 18 new and 22 traditional large high schools.

Student Achievement. The analyses of student achievement on state assessments were conducted with two district samples. The first sample consisted of 11 districts that were in the third or fourth year of their work with the foundation by 2004-05, had multiple foundation-supported schools in operation in 2004-05, and had one or more schools

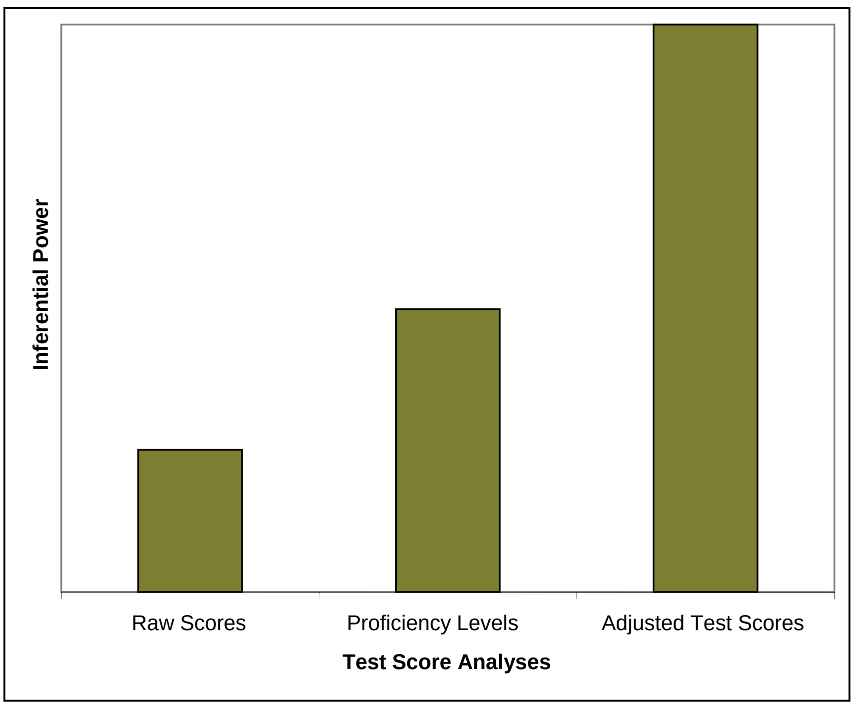
[8] In our discussion of demographic data, we increased the sample sizes by adding two other districts—Sacramento and San Diego. These districts were also ones with which we were familiar, and, although they did not have the data for all of the other analyses available, they did have demographic data available for 2002-03 and 2003-04.

included in our site visit or survey samples.[9] Five additional districts that fit these criteria but did not publicly report average achievement test data at the school level were not included in our analyses.[10] The second district sample used in this section consisted of seven large school districts that provided data used in earlier individual district-specific analyses of achievement trends for different types of high schools.[11] These districts, like the districts in the first sample, had schools in their third or fourth year of work with the foundation by 2004-05. They also were able to provide us with student-level data so that we could conduct analyses controlling for the variation in prior achievement levels of students attending different types of schools (Rhodes et al., 2005). We chose to conduct a separate analysis of this second group of districts because, as discussed below, the second set is exclusively urban and thus reflects the shift in the foundation's funding priorities.

Figure 1 shows the levels of statistical analyses used for student achievement data.

Student achievement can be analyzed by referring to adjusted test scores, proficiency rates, or raw test scores (Figure 1).

Figure 1. Levels of Statistical Analysis Used for Student Achievement Data



[9] These districts were Anderson Union, Baltimore, Chicago, Cincinnati, Cotati-Rohnert Park; El Dorado, New York City (Bronx, Brooklyn, and Manhattan), Oakland, Providence, Ravenswood, and Sacramento.
[10] The districts that could not provide the data were Denver, Milwaukee, Oklahoma City, West Clermont, and Worcester.
[11] These districts were Baltimore, Chicago, Cincinnati, New York City (Bronx, Brooklyn, and Manhattan), Oakland, Providence, and Sacramento.

Figure 1 organizes student achievement analyses conducted by the evaluation team from the statistically least powerful level (raw test scores) to the statistically most powerful level (student test scores adjusted for students' demographic background and prior achievement).

In our series of reports on student achievement in individual school districts, we conducted analyses of student achievement at the adjusted test score level. For this current report, however, our goal was to go beyond the achievement studies in scope and to provide additional analyses for a larger sample of districts. Prior student achievement data from many of the additional districts used in this larger sample were not available, so we have focused our discussion on analyses of unadjusted (raw) test scores. Readers interested in adjusted test score analyses should refer to the series of achievement study reports.

Although Figure 1 suggests that adjusted test score analyses support the most powerful examination of achievement, readers should note that in some policy circles proficiency rate analyses are more important. In this latter view, the most important indicator of student success is whether a student reaches the state-mandated proficiency level in a discipline. The gains in student achievement measured by the other test score analyses, although important, are less important than the gains measured by proficiency rate analyses. This report provides some analyses of proficiency rates as well.

In discussing student proficiency rates on achievement tests, we were constrained by the limited availability of school-level data on individual schools' proficiency rates for consecutive years. Because of this limitation, we could draw on only six of the seven large districts for our analysis of foundation schools' proficiency rates relative to their district averages[12] and on only five for our analysis of the foundation schools' gains in proficiency rates vis-à-vis their respective districts.[13]

Student Work Quality. The analyses of student performance were conducted on a subset of our national sample. We took 1,378 pieces of student work in ELA from 20 schools and 1,246 pieces of student work in mathematics.

For school demographic characteristics, attendance rates, 9th-to-10th-grade progression rates, and achievement scores, we performed descriptive analyses comparing these outcomes in the schools supported by the foundation with district averages in each of the school districts. We present 2 years of data in each of these analyses, and the data each year come from the same schools. We also conducted HLM analyses to compare student work quality in 12 foundation-supported new schools with similar outcome measures in 8 nearby traditional large high schools. To compare student attitudes, we conducted HLM analyses in 18 foundation-supported new schools and 22 traditional large high schools (see the technical appendix for further details on analytic methods). Where appropriate, qualitative data from site visits were used to elaborate or provide context to the findings.

[12] These districts were Chicago, Cincinnati, New York, Oakland, Providence, and Sacramento.
[13] These districts were Chicago, Cincinnati, Oakland, Providence, and Sacramento.

A Reader's Guide for Looking at the Student Outcomes Analyses

Discussions of attendance rates, progression rates, attitudes, achievement scores, and student work quality provide several prisms through which to view student progress at foundation-supported schools, and each discussion helps put the student experience of reforms into a useful perspective. No one indicator tells the entire story, and each is important for a full understanding of the impact of a reform effort. Reaching a conclusion about student outcomes based on the different analyses, however, requires careful consideration of the strengths and weaknesses of each form of analysis.

Readers should note that student attendance rates, progression rates, and attitudes are noteworthy early indicators of student success. They each speak to important conditions for student learning, although they don't necessarily measure student mastery of content or gains in conceptual understanding. Student attendance, for instance, is important because students need to be in school to get credits, to gain the benefit of lessons learned in the classroom, and to be able to interact with teachers and students. Analyses of attendance rates provide us with information about the likelihood that students are getting these opportunities, and increases in attendance are a good sign. Attendance rates alone, however, do not give us any sense of what students are learning.

An analysis of progression rates has similar benefits and limitations. It is important to know whether students are moving from one grade to the next on schedule. Presumably, promotion from one grade to the next speaks to mastery of at least a minimal level of proficiency in students' course work. On the other hand, different schools have different promotion standards, and progression rates do not reveal whether students are meeting minimal standards or excelling.

Student attitudes are another important early indicator. It is unlikely that students will learn much if they express little interest in their course work, have low educational aspirations, or have a low concept of their own academic ability. Any changes in engagement that we can measure speak to students' predisposition to learn, and increases in this area are a good sign.

Student achievement on standardized assessments and student work products are ultimately the most powerful indicators of student learning in our analyses. Each is a different measure of the degree to which students have mastered the curriculum. Student work data describe the extent to which students master course content. Standardized achievement scores describe mastery of a broader range of knowledge and skills, but there are undoubtedly differences between schools in the extent to which schools' learning objectives align with standardized tests. The strength of one measure relative to the other is a matter of debate, and readers would be well advised to consider both of them.

In truth, readers would be well advised to consider all of the analyses of student outcomes discussed in this section. The foundation-supported reforms touch students in many ways, and each of the indicators discussed here contributes to a composite picture of this impact that no one indicator alone could provide.

Discussion of the Findings

Before describing student outcomes, it is important to examine the demographics of students attending foundation-supported schools. The foundation's Education Program targets students from disadvantaged and underserved backgrounds, and the foundation expressly states that it hopes to serve low-income African-American and Hispanic/Latino students (Bill & Melinda Gates Foundation, 2006). This effort is consistent with observations made by the foundation's founders, Bill and Melinda Gates, who frequently have noted the inequities that face disadvantaged students in America's urban schools.[14]

[14] See, for example, Bill & Melinda Gates Foundation (2005b).

To determine whether the schools in our sample were in fact serving the populations that the foundation intended, we examined the demographic characteristics of foundation-supported and other schools in their districts. Our sample for these analyses consisted of nine large urban districts.[15] These districts each provided data on the demographic makeup of foundation-supported schools within the district, as well as district average enrollments of students from each of our designated demographic groups.

Using this information, we compared enrollments at the foundation-supported schools with the respective averages from their entire districts, focusing specifically on the following demographic groups:

- Students with low socioeconomic status, as measured by eligibility for free or reduced-price lunch
- Racial/ethnic minority students
- Students receiving special education services, as measured by those with an individualized education program (IEP)
- English language learner (ELL) students

We report the percentage of schools within each category that had enrollments of the targeted population that were above their own districts' average, so that we can make comparisons across districts. Districts generally vary in the way they measure various outcomes. For instance, because schools' respective districts may measure rates differently, direct comparisons of schools' enrollment rates or progression rates are meaningless when the compared schools are in different districts. On the other hand, in comparing the enrollment or progression rates of different schools with their respective districts' averages, the districts' different methods of measuring the rates do not matter because we are comparing each school's performance with that of other schools in its district. We can therefore use this measure to make the comparisons across districts without having to worry about differences in the districts' methods of measurement.

Although the foundation does not specifically target students requiring special education services or ELL students, we recognized that both groups are among the nation's most disadvantaged students and that, as such, serving them falls within the general spirit of the foundation's initiative. Including these groups in our analysis also allows us to determine whether the foundation-supported schools are enrolling smaller percentages of students from demographic groups that generally have low achievement test scores and are harder to teach.

Serving Targeted Students

Two thirds of new schools and almost 80% of redesigned schools exceeded their district averages for enrollment of students eligible for free or reduced-price lunch and for enrollment of students from minority backgrounds.

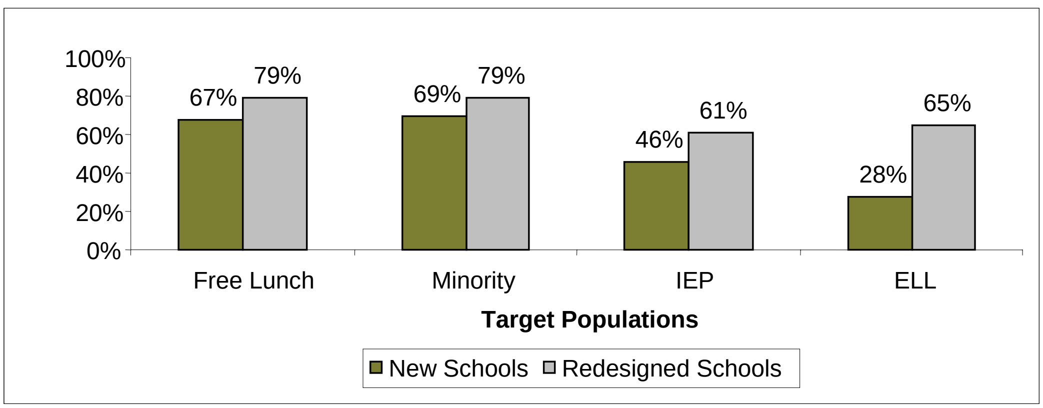
[15] These districts were Baltimore, Chicago, Cincinnati, Los Angeles, New York, Oakland, Providence, Sacramento, and San Diego. Three of the districts had fewer than five new schools or fewer than five redesigned schools; findings from these districts may be less reliable.

Nationally, about 32% of high school students have minority backgrounds (National Center for Education Statistics, 2005), and 13% have IEPs. In the nine districts in our sample, however, students from minority backgrounds on average made up 62% of their district enrollments in 2000, and in the 2000-01 school year, an average of 13% of the students in these districts had IEPs (National Center for Education Statistics, 2005). In the nation's largest urban districts—a grouping that includes these nine districts—about 29% of high school students are eligible for free or reduced-price lunch (National Center for Education Statistics, 2005). Just by making grants in these districts, the foundation is targeting the population it wants to serve.

A closer look at the data suggests that even within these districts, the foundation-supported new and redesigned schools stood out in terms of serving the foundation's targeted population. Two thirds (67%) of the new schools and almost 80% of the redesigned schools exceeded their district averages for enrollment of students eligible for free or reduced-price lunch. Similar results were found for the enrollment of minority students (see Figure 2).

In 2004, more than 60% of new and redesigned schools had enrollments of minority students and students eligible for free or reduced-price lunch that exceeded district averages, but only redesigned schools generally had higher-than-district-average enrollments of ELL students and students using IEPs (Figure 2).

Figure 2. Percentage of Schools With Above-District Average Enrollments of Targeted Populations, by School Type, 2004



Sample. Total number of new schools = 57; data were unavailable for two new schools in one district. IEP data were unavailable in one district, and ELL data were unavailable in another. Total number of redesigned schools = 28; ELL data were unavailable for one district.
Data sources. Extant data from nine districts

Conversely, not all foundation-supported schools were serving all disadvantaged student populations to the same degree. Whereas more than 60% of foundation-sponsored redesigned schools had enrollments of students with IEPs and enrollments of ELL students that exceeded their respective district averages, fewer than 50% of the foundation-sponsored new schools did so. The percentage of new schools with enrollments of ELL students above the districts' averages was particularly low (28%). The percentage (46%)

of new schools with enrollments of students with IEPs that exceeded their district averages was not a negligible figure, but it was noticeably smaller than the percentages of new schools that had above-district-average enrollments of minority students and students eligible for free or reduced-price lunch. More than 65% of the new schools had above-district-average enrollments of these latter groups of students.

We cannot conclude with certainty why the new schools were below their district averages in enrolling ELL students. In some instances, these students may have been missed when the new schools engaged in student recruitment. We also suspect, however, that, because of their small size and limited number of faculty members, newer schools in some cases simply had a hard time meeting the needs of ELL students adequately. These students might have left the schools or might have been directed to attend other schools that had better capacities to serve special student subgroups. Recognizing the difficulty it had in serving ELL students, for instance, one new school in our sample chose to eliminate its program for these students and to create a separate school for them. This specialized school was expected to focus only on these students' learning needs.

The finding that nearly 50% of the new schools had enrollments of students with IEPs that exceeded district averages indicates that although the new schools were not enrolling these students to the same degree that they were enrolling students from minority backgrounds or those eligible for free or reduced-price lunch, their enrollments of the these students were still substantial. This finding is noteworthy, particularly when we consider that teachers at several new schools commented on the difficulty they experienced serving students with special educational needs. In one instance, the special education teacher at a school worked with students for no more than half of the school day. The district was supposed to provide additional services but had been unable to provide staff. Although we do not have confirming data on the point, one can imagine that these scenarios might encourage families of students with special education needs to look to other schools for a more supportive environment.

Attendance

In the districts that we studied, almost 80% of the foundation-supported new schools had attendance rates that were higher than their district averages in 2002-03, and this percentage increased to 85% in 2003-04. The attendance rates for most foundation-supported redesigned schools, however, were below district averages.

Attendance rates are an indicator of the degree to which a school has been able to foster a culture in which students maximize opportunities to learn by attending classes regularly, and they are an important early indicator for students' success in school. Our analysis of attendance focused on six school districts.[16] We counted the total number of

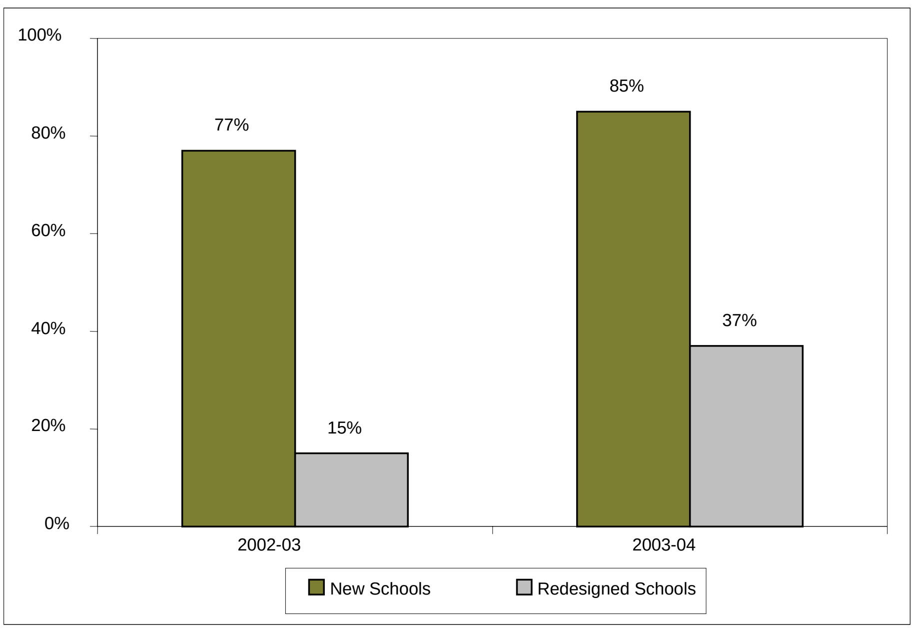
[16] These districts were Baltimore, Chicago, Cincinnati, New York, Oakland, and Providence. These six districts, in addition to hosting several foundation-supported schools, were also the only ones we worked with that were able to provide us with data for both the 2002-03 and 2003-04 school years.

new and redesigned schools across the districts and determined the percentage of schools within each type with average daily attendance rates that exceeded their respective district averages. We report the average daily attendance rates at both types of schools for the 2002-03 and 2003-04 school years. The results represent findings from the same set of schools in both years.

As shown in Figure 3, 77% of the new schools in 2002-03 and 85% of these schools in 2003-04 had attendance rates that were higher than their respective district averages. The comparatively stronger performance of new schools undoubtedly reflects the continuing enthusiasm—or discipline—that can take hold in new learning communities. Partly because of their small size, new schools could set attendance policies that were reinforced by the closer contact and more personalized relationships between teachers and students. The new schools might also have benefited from the fact that they were schools of choice for many students, who opted to attend the institutions, as opposed to being assigned to them.

In 2002-03 and 2003-04, substantial majorities of new schools had attendance rates above their district averages, while most redesigned schools did not (Figure 3).

Figure 3. Percentage of Schools With Average Attendance Rates Higher Than Their District Averages, by School Type, 2002-03 and 2003-04



Sample. Number of new schools = 26; number of redesigned schools = 27
Data source. Extant data from six districts

By contrast, redesigned schools as a group fared poorly when their attendance rates were compared with their respective district averages. Indeed, only 15% of the redesigned schools had attendance rates above their respective district averages in 2002-03. The redesigned schools' attendance rates did improve, however, relative to their district averages in 2003-04. Although we should not draw causal conclusions based on the change from 1 year to the next, we have seen several possible explanations for this improvement in our fieldwork. A teacher in a redesigned school suggested, for instance, that attendance had improved at her school because students had a smaller group of teachers. She referred to her small learning community that the schools created by saying "we all set the standards that we want students to meet," and these consistent standards have helped "students realize that they have to produce, that they need to be in class every single day, and that they need to be working."[17] Other strategies that we have observed at redesigned schools include (a) the implementation of Friday afternoon detention for missing class and (b) schools' contacting parents when students do not arrive at school in the morning. One school even has a staff member charged with going to students' homes and bringing the students to school. As one school leader noted, "We are really setting the expectation that people have to attend. We need to continue to work with students and parents to get them to do that."

Progression Rates

In 2002-03 and 2003-04, the 9th-to-10th-grade progression rates at new high schools compared well with the schools' respective district averages in the four districts for which data were available. All of the new high schools in our sample had 9th-to-10th-grade progression rates above their district averages in 2002-03, and close to 60% did so in 2003-04.

Reducing student truancy is merely a first step in improving the culture of learning in school. Beyond ensuring student attendance, school leaders and communities need to be concerned with whether students are learning enough to make progress from one grade level to the next and whether the students are ultimately making enough progress to graduate. Consequently, as another indicator of the impact that foundation-supported schools may be having in fostering success for all students, we examined student progression rates in foundation-sponsored schools and compared these with the schools' respective district averages.

Many of the schools in our study are too new to have any data on graduation rates. However, in a recent report, the University of Chicago's Consortium on Chicago School Research found that an early indicator of students' being on track to graduate from high school is their ability to meet the requirements to progress from 9th to 10th grade (Allensworth & Easton, 2005). Therefore, to estimate the future graduation rates in new schools, we examined the percentage of students who met their district's requirements for promotion from 9th to 10th grade.[18]

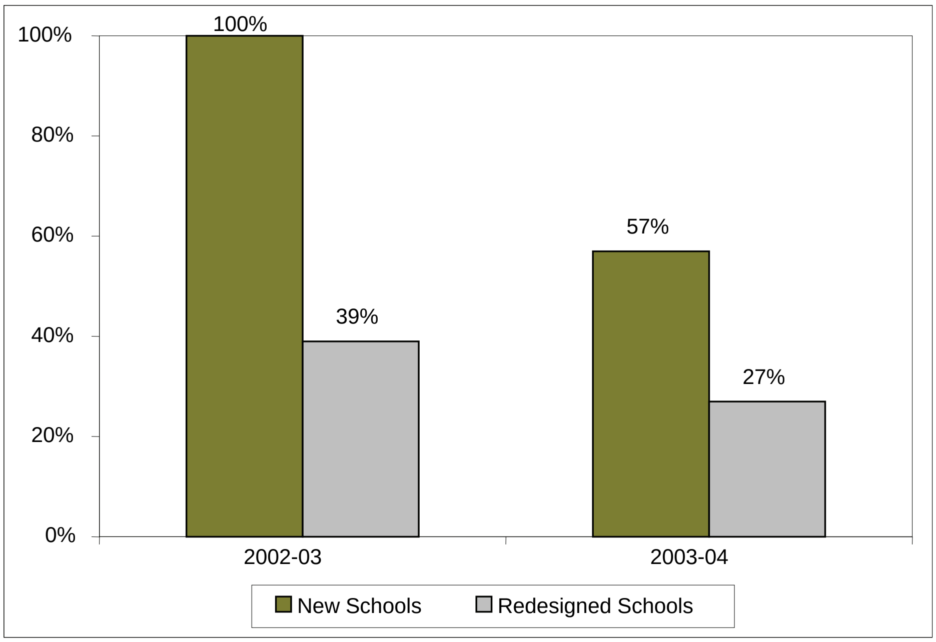
[17] Our limited number of schools in each district did not allow a model to adjust for characteristics that may systematically affect attendance rates.

[18] Progression rates were calculated by dividing the number of students who had moved on to 10th grade by the number of students in 9th grade during the previous year. Students who transferred to another school or whose status the next year was unknown were deleted from the ninth-grade population.

Our analysis of progression rates is limited to the four districts that were able to supply usable data on student progression.[19] Figure 4 displays the percentages of schools within each school type that had 9th-to-10th-grade progression rates above the schools' respective district averages. As with attendance rates, we report the progression rates for both the 2002-03 and 2003-04 school years, and the results are taken from the same schools in both years.

In 2002-03 and 2003-04, substantial majorities of new schools had average 9th-to-10th-grade progression rates higher than their district averages, while most redesigned schools did not (Figure 4).

Figure 4. Percentage of Schools With Average 9th-to-10th-Grade Progression Rates Higher Than Their District Averages, by School Type, 2002-03 and 2003-04



Sample. Total number of schools = 25 (new, 7; redesigned, 18)
Data sources. Extant data from four districts

As Figure 4 indicates, 100% of the new schools had 9th-to-10th-grade progression rates that were above their district averages for that year. The personalization of relationships between teachers and students at new schools appears to be an important factor contributing to new schools' noteworthy progression rates. In a discussion at one new school, for instance, students made similar observations about how teachers helped them progress in school. One student commented, "Our teachers are going to stay on you and ride your back about doing homework. They really want you to pass, and they stay on you about doing your work because they don't want to see any student left behind." A

[19] These districts were Baltimore, Chicago, Cincinnati, and Providence. Unfortunately, the other three districts (Los Angeles, New York City, and Oakland) were either unable or unwilling to provide us with the data.

second student added that his advisors "call [his] mom all the time. They tell if you are doing well or bad. They tell you if you are in danger of not passing."

In 2003-04, the progression rates at 57% of the new schools still exceeded those of their districts. However, this percentage is a notable decrease from the 100% from the previous year. The decrease between 2002-03 and 2003-04 is consistent with dips we have seen in other outcomes as new schools move beyond their initial year. For example, a teacher at one maturing new school explained that the school experienced "growing pains" after it first opened, and staff members were still figuring out how to run the school effectively. The school was originally modeled after another high school, but staff learned that they had to "tweak the model" a bit to create more structure for the students. The school is creating criteria to base grade progression on mastery, not just effort, and is realizing that some students need to be held back after the first year.

Similarly, some new schools may be increasing the rigor of their course work as they mature—a goal that is consistent with the hopes of the foundation. Under this scenario, as students get pushed into more challenging courses, the weaker students in the group may face greater academic difficulties as their weak preparation surfaces. These students may then grow discouraged and disengage to the point where they do not progress from 9th to 10th grade.[20]

Progression rates at redesigned schools were probably low for many of the same reasons that the rates were low at new schools. Additionally, we should point out that many of the redesigned schools were really the offspring of parent schools with some of the weaker student performances in their respective districts. Because the redesigned schools were generally working with the same student bodies as their parent schools had, we might expect their progression rates to reflect this historical weakness when compared with district averages.

Finally, we should note that the progression rates remained stable from one year to the next in most of the schools in our sample. There were two new schools that experienced instability in their progression rates. At one school, the drop was 17 percentage points. Another school, however, had a difficult first 2 years, and its progression rate dropped precipitously—more than 60 percentage points—from the first year to the second. Our comparison of progression rates with district averages should be considered with this instability in mind.[21]

Student Attitudes

In 2004-05, students in foundation-supported new high schools in our sample had aspirations for their futures that were higher than those of students in other district high schools. Students in the new schools reported plans to graduate from high school and

[20] Our review of teacher assignments, however, does not support this hypothesis. In our analyses, we found that the rigor of teachers' assignments actually declined over time in the new schools that we studied.

[21] We considered triangulating this finding with some of our student attitude indicators to see whether we found parallel patterns in the second and third years. Unfortunately, our surveys collecting data on student attitudes were administered in off-years from the progression rate data, so triangulation was not possible.

apply to college at rates that were higher than those of students in other schools. Students in the new schools also reported greater levels of engagement in terms of persistence and interest in their studies.

The research literature (Akey, 2006; Bransford, Brown, & Cocking, 2000; Newmann, 1992) suggests that both student engagement and students' perceptions about their academic competence have a positive influence on students' high school achievement. Our own analyses suggest that, relative to their district averages, students' engagement and aspirations are relatively high in schools with higher ELA and mathematics achievement test scores. Consequently, in addition to analyses of academic student outcomes—that is, outcomes that directly measure students' academic performance (e.g., test scores and student work quality)—we also examined selected measures of students' attitudes about schooling, using responses from the student survey. Below is a brief description of these measures:

- **Engagement—Interest:** A measure of how often students asked questions in class or contributed to class discussions, met with teachers to talk about schoolwork, talked to family about schoolwork, asked friends for advice about schoolwork, and worked with classmates outside of class or school on schoolwork
- **Engagement—Persistence:** A measure of how often students made extra effort on challenging assignments, got help with difficult homework, and tried to do well on schoolwork even if they thought it was too hard or not interesting
- **Academic self-concept:** A measure of the degree to which students felt they were good at reading, writing, learning mathematics, getting help, and working with others
- **Educational aspirations:** A measure of how far students planned to go in schooling, ranging from dropping out of high school to postcollege education

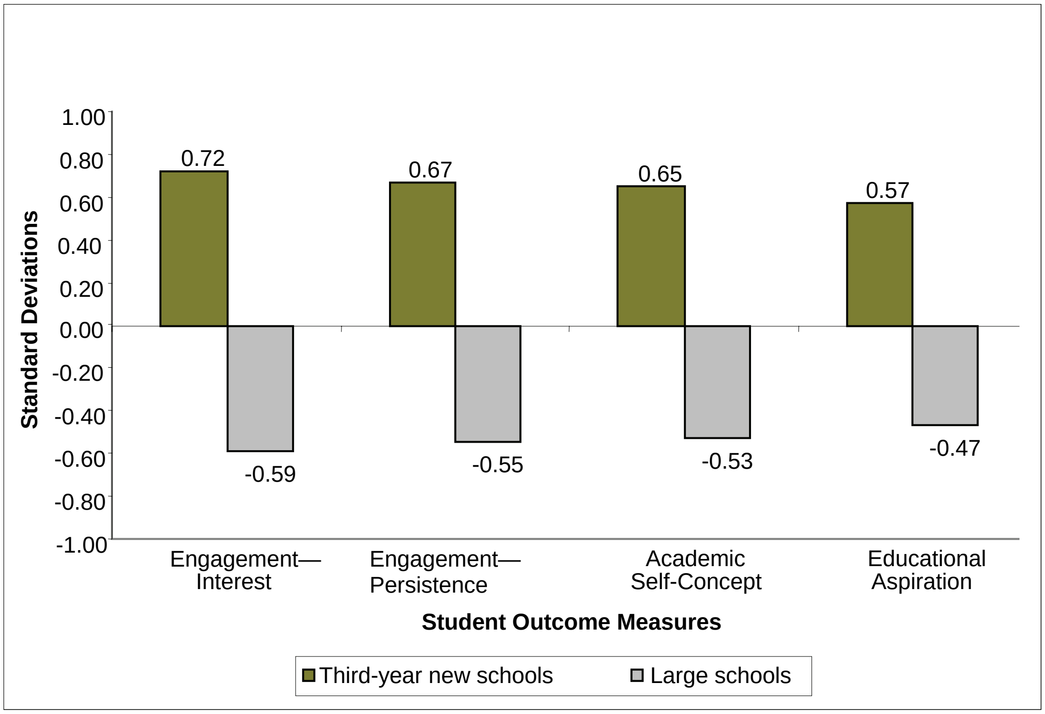
Our analyses focused on 18 third-year new high schools, which were the most mature new high schools in our sample. Results from these schools were compared with 22 traditional large high schools.[22] The results of the analyses are presented in Figure 5.[23]

[22] In this instance, the term *large schools* includes the schools that were targeted for redesign as well as the set of comparison schools.

[23] We have used adjusted rates in this analysis. Adjusted rates take out differences that can be explained by the characteristics of students entering each school, including differences in race/ethnicity, gender, and economic status, as well as the schools' overall minority and socioeconomic status composition.

Students in foundation-supported third-year new high schools demonstrated higher levels of interest and persistence, more positive academic self-concepts, and higher educational aspirations than students from other large high schools (Figure 5).

Figure 5. Differences in Student Attitudes Between Third-Year New High Schools and Large Schools



Sample. The 18 third-year new schools included 8 new schools that opened in 2002 and 10 new schools that opened in 2003. The 22 large schools included 8 preredesign schools that planned for redesign in 2002, 2 preredesign schools that planned for redesign in 2003, 2 preredesign schools that planned for redesign in 2004, 3 preredesign schools that planned for redesign in 2005, and 7 comparison schools surveyed in 2004 (3) and 2005 (4).
Data sources. AIR/SRI student surveys

Similar to what we found about first-year new high schools in last year's student outcomes report (Rhodes et al., 2005), students in third-year new high schools demonstrated levels of interest and persistence, more positive academic self-concept, and higher educational aspirations that were significantly higher than those of students in large high schools ($p < .01$), even after adjusting for student demographic characteristics and school demographic composition. (See the technical appendix for details on analytic methods.)

In a separate analysis, we compared the levels of engagement of students in third-year new schools with those of students in a set of model schools that, according to the foundation, had high levels of implementation of the effective-school attributes that were introduced in chapter 1 (AIR/SRI, 2004). Although students in third-year new high schools did not reach the levels of interest, persistence, academic self-concept, or educational aspirations reported by students in model schools, the differences between the two groups were not statistically significant.

We also examined change over time in student attitudes. Overall, student engagement and academic self-concept remained stable over time in new high schools. New high schools, however, experienced a significant drop ($p < .05$) in the level of students' educational aspirations from their first year to their second year, but aspirations rebounded in the third year to a level comparable to that of the first year. In redesigned schools, significant improvement ($p < .05$) occurred in student engagement, but not in students' academic self-concept and educational aspirations, when attitudes in the parent schools in the year prior to redesign are compared with those of students during the second year of redesign. Detailed results from analysis of these changes are provided in Tables A-2 through A-5 in the technical appendix.

Achievement Scores

In the studied districts, the average achievement test scores reported by most new and redesigned schools were below district averages on standardized tests in ELA and mathematics. The percentages of new schools with average achievement test scores above district averages in the two subjects were higher than the percentages of redesigned schools with above-average scores. In a very small sample of districts and schools, more than 50% of the new and redesigned schools made gains in ELA and mathematics proficiency relative to their respective districts.

Although academic accountability and assessment movements have been gaining momentum since the mid-1990s, the passage of the No Child Left Behind (NCLB) Act of 2001 heightened high schools' focus on standards-based curriculum and testing. NCLB directs states to establish statewide academic skill proficiency levels for students in grades 3 through 12 and to refine, adapt, or create new assessment instruments that help gauge whether students are meeting these proficiency levels. High schools that fail to meet NCLB's standard for "adequate yearly progress" in increasing student proficiency levels face significant penalties. Within this highly charged environment of accountability, a wealth of new data regarding student performance is now more widely available for analysis, and we have chosen to include some of these data in our evaluation.

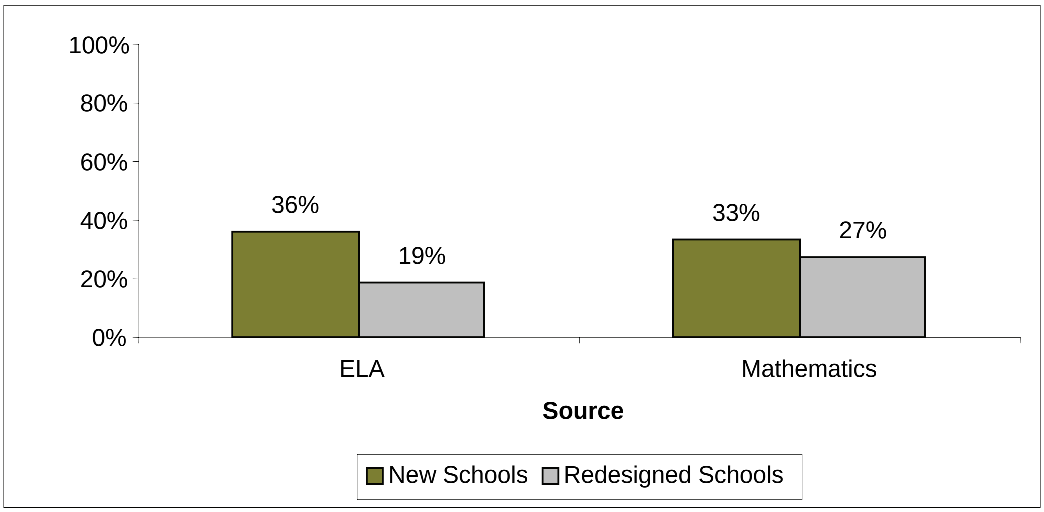
As with attendance and progression rates, we performed within-district and across-district analyses of student performances on statewide assessments.[24] Because we were particularly interested in the work of schools with some tenure with the foundation, and because many high schools administer standardized tests in 10th but not 9th grade, this particular analysis included only schools that were in their second or later years of operation or redesign. For consistency, we limited our analyses to 10th-grade academic achievement.

[24] Making comparisons across states is not meaningful in a discussion about assessment data because instruments and proficiency levels vary widely across states.

Figure 6 presents the percentages of schools by type in our larger district sample whose students' average scores were above their respective district mean scores in ELA and mathematics. The scores presented have not been adjusted to reflect the students' various demographic characteristics or prior academic achievements.

In 2003-04, fewer than 40% of foundation-sponsored new and redesigned schools had average ELA scores that were higher than their district averages, and fewer than 30% had math scores that were higher than district averages (Figure 6).

Figure 6. Percentage of Schools in Operation for at Least 2 Years With Average Achievement Scores Above Their Respective District Means in ELA and Mathematics, by School Type, 2003-04



Sample. Total number of schools = 52 (new, 25; redesigned for ELA, 16; redesigned for mathematics, 11)
Data sources. Extant data from 11 districts

Figure 6 indicates that 36% of the new schools in our sample had average ELA scores that were higher than their respective district averages, and 33% of the new schools had mathematics scores that were higher than district averages. The figures from the redesigned schools were not as strong as those from the new schools. Only 19% of the redesigned schools had average ELA scores that were above their respective district averages, and only 27% had mathematics scores that were above district averages.

The weaker performance of the redesigned schools compared with the new schools may reflect differences in the makeup of the schools' respective student bodies. A new school often benefits from having a student population that has self-selected to attend the school. The families and students who have chosen to leave previous schools in order to attend new ones may very well be more motivated toward academic achievement. By contrast, redesigned schools are on the same campus as the parent institution from which they were created. The students at the redesigned schools may be attending the schools by default; that is, they may not have had to make a special effort to attend but instead have returned to the same campus that they had previously attended. They may also be assigned to a small learning community if they themselves have not taken the initiative to select one. The lower performance at the redesigned schools may therefore reflect a lower level of academic motivation in the student body compared with that of the new schools.

Overall, foundation-supported new and redesigned schools in the large urban districts had lower percentages of students scoring above the district averages on statewide ELA assessments than what was found in our analysis of the larger, more heterogeneous district sample.[25]

Over time, the foundation's funding priorities have changed. In recent years the foundation has chosen to give particular attention to large urban districts where schools have had particularly weak histories of student achievement. These districts have been the topic of separate reports on their respective levels of student achievement, and they have been asked to provide us with data that other districts in our sample have not. When we review the achievement scores from the new and redesigned schools in these large urban districts separately, the results are not as encouraging. Only 29% of the new schools in the large urban districts had average achievement scores in ELA that were higher than district averages. For the redesigned schools, the figure was even lower; only 19% of the redesigned schools in these urban districts had ELA averages above those of their districts. The performances of both types of schools in mathematics were a bit stronger, as 33% of the new schools and 27% of the redesigned schools were above the averages in these urban districts.

The weak performance of students at both types of schools on the assessments may be attributable largely to the weak academic foundation that many of the students had before they entered the high schools. In our summary of our achievement studies, we remark that many of the students in foundation-supported high schools in large cities had histories of poor academic performance and that the students' average scores on eighth-grade proficiency tests were typically well below the proficiency levels announced by their districts.[26] This weak preparation appears to be a lingering obstacle for the students.

In both of our district samples, weak academic foundations were a burden for students entering the two types of schools. A teacher at one high school, for instance, commented that his students' skills were "atrocious," adding that he had "16-year-olds with second-grade reading levels." A teacher in another district reported that, on average, students came to his school with a fifth-grade reading level and pointed out that reading problems can also undermine mathematics achievement. "If they can't read, they can't do math word problems," he commented.

In mathematics, the fact that concepts build sequentially and are interdependent appeared to exacerbate the challenge posed by students' entering several grade levels

[25] These districts were Anderson Union, Baltimore, Chicago, Cincinnati, Cotati-Rohnert Park; El Dorado, New York City (Bronx, Brooklyn, and Manhattan), Oakland, Providence, Ravenswood, and Sacramento.
[26] See the individual achievement studies in AIR/SRI (2005).

behind standards. One teacher noted, "It's really hard to catch up with math when you have this great gap" of five or six grade levels, which some of her students have had. A shortage of qualified mathematics teachers may also contribute to the students' relatively poor mathematics performance. On site visits, we heard a number of stories about the difficulties that school leaders have had in hiring good mathematics teachers.

A comparison of average scores on assessments with district averages tells only part of the story of student achievement. Ultimately, we must be concerned with the percentages of students who reach their respective state standards for proficiency in ELA and in mathematics. If the percentages of students reaching proficiency levels do not improve, then the improvements in test scores, although commendable, represent a limited accomplishment.

Limited data suggest that foundation-supported schools may be making gains in the percentages of their students meeting local proficiency rates compared with other schools in their districts.

Arguably, the most important indicator of the progress made by foundation-supported schools in the area of student achievement is how their proficiency rates compare with the average proficiency rates of schools in their districts. If the proficiency rates are lower, another question is whether the schools have made gains in their proficiency rates vis-à-vis their districts. If gains in proficiency in the foundation-supported schools are greater than the gains in the districts overall, then students in the foundation-supported schools are making noteworthy progress.

In 2003-04, a substantial percentage (57%) of the new schools in our sample had ELA proficiency rates that were higher than their respective district averages, but a much smaller percentage (24%) had higher mathematics proficiency rates. The percentage of redesigned schools with higher-than-district-average proficiency rates was comparable to that of new schools in mathematics but much smaller in ELA. Only 27% of the redesigned schools had proficiency rates that exceeded their district averages. On the basis of these data, the foundation-supported new schools compare well on student achievement in ELA, but neither new schools nor redesigned schools compare well in mathematics. Redesigned schools also compare poorly in ELA.

Over time, however, students in foundation-supported schools—while still reaching proficiency at the same or lower rates than their peers—may be making greater gains. For a very small sample of foundation-supported schools, we compared the gains over time in the percentage of proficient students with the average gains for the district. Table 4 shows the percentages of foundation-supported schools that had greater gains than their district.[27]

[27] Data in this table are limited to those districts from which we had proficiency rate data for individual schools in consecutive years, as well as district-level proficiency rates for the same years. The districts for which these data were available are Chicago, Cincinnati, Oakland, and Providence. To enlarge our sample, we added Sacramento, but data were not available for Baltimore, Los Angeles, and New York.

Table 4. Percentages of Foundation-Supported Schools With Higher Proficiency Rate Gains Than Their Districts

	2002-03 to 2003-04	2003-04 to 2004-05
ELA		
Redesigned Schools	66%	57%
New Schools	75%	66%
Mathematics		
Redesigned Schools	33%	57%
New Schools	80%	33%

Sample. Number of redesigned schools = 10 (3 for 2002-03 to 2003-04 and 7 for 2003-04 to 2004-05); number of new schools = 8 (5 for 2002-03 to 2003-04 and 3 for 2003-04 to 2004-05)
Data sources. Foundation Strategy Group (2005), AIR/SRI achievement studies on five districts (AIR/SRI, 2005), and Standard and Poor's School Evaluation Services Web site: http://www.schoolmatters.com

Although the available data are extremely limited, they present an encouraging portrait of the foundation-supported schools' gains in ELA student achievement. Between the 2002-03 and 2003-04 school years, 75% of the new schools made gains in the percentages of students achieving ELA proficiency, and in the following year 66% made gains. Redesigned schools also had encouraging ELA results, with 66% making gains in the first year and 57% making gains in the second year.

The story in math scores, however, is mixed. Among new schools, 80% of the schools for which data were available made gains in math proficiency rates between 2002-03 and 2003-04 but only 33% between 2003-04 and 2004-05. The redesigned schools had similarly inconsistent gains. Thirty-three percent made gains in mathematics between 2002-03 and 2003-04, while 57% made gains the following year.

In making these comparisons in these five districts and in the districts overall, one should keep in mind that many foundation-supported schools are working with some of their districts' historically lowest-achieving students. Consequently, we might expect the percentages of students in these schools who reach proficiency to be smaller than the percentages from the districts overall. Data suggesting that the gap in student proficiency rates between these schools and their districts might be decreasing may augur well for the future. On the other hand, readers should also bear in mind that, given the size of the samples here, events at a small number of schools can make percentages shift substantially. Consequently, these findings should not be considered conclusive.

Student Work Quality

On the basis of student work collected in 2003-04 and 2004-05, students in foundation-supported new high schools were engaged in higher-quality work in ELA classes than were students in traditional high schools without foundation ties. The quality of students' work in mathematics classes in the foundation-supported new high schools was similar to that of students in other schools.

Many critics of the high-stakes testing movement maintain that a test score does not adequately capture a student's intellectual development or demonstrate what a student knows and can do. Instead, many of these critics claim that student learning is more appropriately assessed by a review of student work samples. In this section, we analyze student work products as another way to measure academic achievement. We specifically look at classwork done by students in order to assess whether students in foundation-supported schools were producing higher-quality work than students in other schools.

Overall Student Work Quality

We sought to capture the quality of student work by the extent to which student work products demonstrated the following:

- Construction of new knowledge using existing knowledge (in ELA)
- Effective communication and accurate use of language and conventions (in ELA)
- Procedural knowledge (in mathematics)
- Deep conceptual understanding of important content (in mathematics)
- Reasoning and problem-solving facility (in mathematics)

Our analysis of student work focused only on foundation-supported new schools and a set of comparison schools. We are not yet able to analyze student work in the redesigned schools, because data are currently being collected. Results from preredesign and redesigned schools will be presented next year.

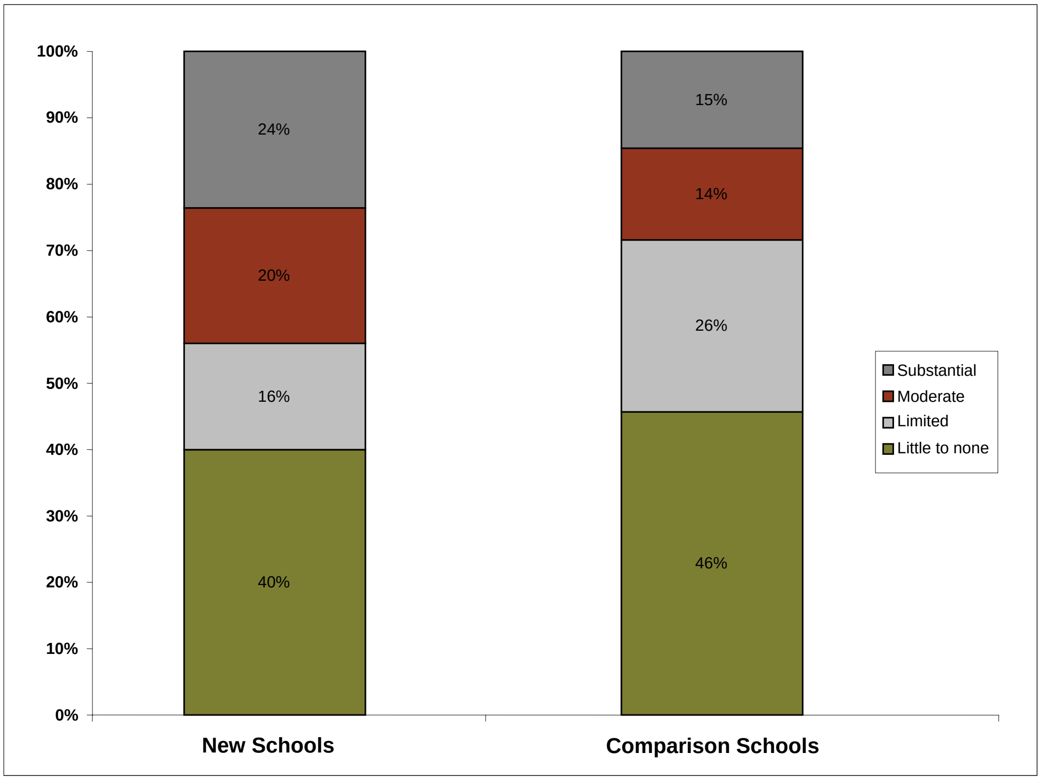
Using Many-Facet Rasch Measurement, a psychometric measurement model, we derived an overall score measuring the level of quality for each piece of student work. The master teachers scoring the work were not told which type of school (new or comparison) the student attended. (For further information on the subject-specific criteria for grading the quality of student work, see the technical appendix.)

The results of our analyses of student work quality in ELA are encouraging for the new schools, in that work quality was significantly higher in new schools in our sample than in comparison schools. Overall, we were able to classify student work in ELA along the following gradations of quality: substantial, moderate, limited, little to none.

At the upper end of our gradations of work quality, student work at new schools generally exceeded that of comparison schools in quality (see Figure 7). In the new schools, 44% of the work was of moderate quality or better; in the comparison schools, only 29% was rated as highly. However, more than 50% of the work in these schools fell within the limited and little to none quality ranges, with 40% of the student work at new schools rated as having little to no quality. At large comparison schools, 46% of the work was rated as having limited or little to no quality.

ELA student work in new schools generally exceeded that of comparison schools in quality (Figure 7).

Figure 7. ELA Student Work Quality



Sample. Number of assignments from new schools = 634; number of assignments from comparison schools = 292
Data sources. Number of individual schools providing student work = 20 (new schools, 12; traditional large schools, 8)

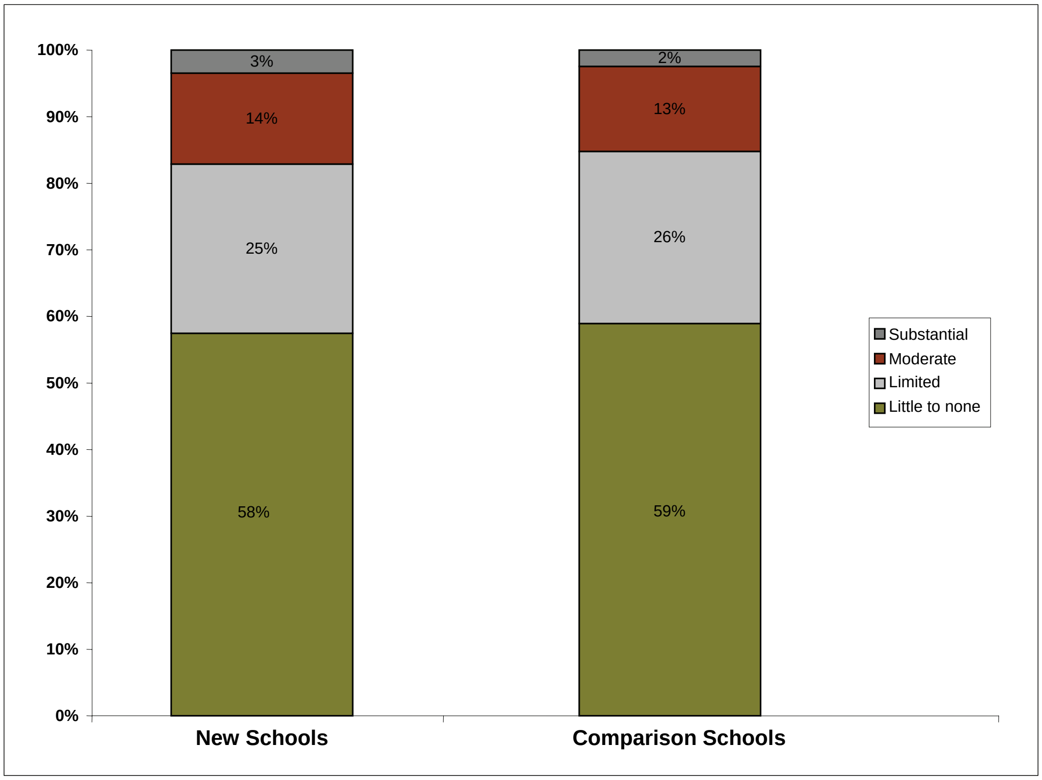
The foundation's emphasis on the relevance of school assignments may be one explanation for the higher quality of student work in ELA at the new schools. In chapter 3, we note that the ELA assignments that teachers in new schools have been giving to students have generally been more relevant to student interests and real-world contexts than those given in the comparison schools. We note further that relevant ELA assignments generally lead to higher student outcomes.

The ratings of student work quality in mathematics were much less positive for new schools. We assigned students' classwork from new and comparison schools to gradations along the same categories described above for ELA. Our data are presented in Figure 8.[28]

[28] These analyses did not use HLM, nor did they control for variables in students' backgrounds. When we used HLM, results showed that the ELA coefficient of new schools was positive and statistically significant ($p < .01$), and the math coefficient was not statistically significant ($p = .98$).

Almost 60% of the student work in mathematics was rated as being of limited to no quality in both new and comparison schools (Figure 8).

Figure 8. Mathematics Student Work Quality



Sample. Number of assignments from new schools = 519; number of assignments from comparison schools = 300
Data sources. Student work provided by individual schools. For greater detail, see the technical appendix.

As Figure 8 shows, we found little to no quality (i.e., the work generally did not show conceptual understanding of important mathematics, procedural knowledge, problem solving and reasoning, or effective communication) in almost 60% of the student work in mathematics at both new schools and comparison schools.[29]

There are several possible explanations for the low quality of student mathematics work in these schools. One is the ongoing difficulty that many schools are having in finding certified or highly qualified mathematics instructors. There are shortages of certified mathematics instructors in high-poverty districts all over the country. We noted in our discussion of student scores on mathematics assessments that many of the new schools

[29] A caveat should be mentioned in interpreting these classifications of student work. This analysis is based solely on the total number of assignments and student work turned in. There are no controls for any one school's proportion of the total output of student work product. If a school with a weak mathematics program has turned in a disproportionate number of student work products, then that school will skew our results.

in our sample have had difficulties in finding such teachers, and many of the comparison schools may share this problem.

Curricular and pedagogical limitations may also be part of the explanation for poor performance in mathematics. The number of teachers in our sample who have access to mathematics curricula that dovetail nicely with our evaluative criteria (i.e., curricula that call for work products that allow students to construct knowledge, demonstrate a deep conceptual understanding of important content, etc.) may be limited.

An example may help illustrate this last point. In traditional mathematics teaching, students are often given assignments that require them to provide short-answer responses to problems posed in a textbook. To complete these assignments, a student might be challenged, have to work hard, and develop an understanding of mathematics concepts. The limited nature of this assignment, however, precludes the student from demonstrating his/her conceptual understanding. And if the assignment does not ask the student to show his/her work, then the student cannot demonstrate procedural knowledge or effective communication of mathematics concepts. Student work of this sort would receive a low score when our rubric was applied, even though the student submitting the work might have learned far more than the assignment asks him or her to demonstrate.

Finally, one should not underestimate the continuing impact of traditional mathematics pedagogy and testing. Despite many of the schools' efforts to move mathematics instructors toward more problem-based pedagogies, we have learned from site visits that mathematics skills are often taught with traditional recitations and testing. If mathematics is taught in a conventional fashion, then even at the new schools stressing project-based learning in other subjects, there is little reason to expect that students' work will be different from that of students in comparison schools.

Implications for the Initiative

A discussion earlier in this chapter presented the strengths and weaknesses of different indicators of student success. The discussion distinguished between early indicators—attendance rates, progression rates, and student attitudes—and more summative measures, such as student achievement scores and the quality of student work. The review of the early indicators supports cautious optimism about student outcomes in foundation-supported new schools. Outcomes in new schools' attendance and progression rates are particularly encouraging. Close to 60% of the sampled new schools had higher-than-district-average 9th-to-10th-grade progression rates, and more than 80% of them had above-average attendance rates. Student work in ELA at sampled new schools was generally of higher quality than student work from comparison schools; student work in mathematics was similar at both types of sampled schools; and students in third-year new schools generally had higher levels of engagement, more positive academic self-concept, and higher educational aspirations than students in large high schools.

The caveat is that analyses of the more summative indicators in the new schools are less encouraging. Student scores on achievement assessments remain a problem for the new schools in our sample. Fewer than half of the new schools in our larger sample of school districts had average ELA or mathematics scores that exceeded their respective district averages. When we looked separately at results from the large school districts that recently have become the foci of the foundation's efforts, we found even smaller percentages. In the larger districts, the percentage of new schools with average scores above the district averages in either discipline did not exceed 33%. Student work in mathematics in new schools is also a reason to be concerned. In our analysis of student work in mathematics, we found little to no quality—as we have defined it—in nearly 60% of the student work at sampled new schools (and comparison schools). Improvement in the quality of student work in ELA was the one strong exception to this pattern. Whereas 44% of the student work in new schools was found to be of moderate quality or better, only 29% of the work in comparison large schools was rated as highly.

After starting its initiative, the foundation quickly concluded that starting good new schools is easier than profoundly improving existing high schools. Consistent with this conclusion, our analyses confirm that student outcomes in our samples are less positive in the foundation's redesigned schools. Student attendance rates, 9th-to-10th-grade progression rates, and achievement scores at these schools consistently fell below district averages. On the basis of these indicators, the students at the redesigned schools are not meeting with the same preliminary success as the students in the new schools.

There may be some encouraging news with regard to student achievement. In a very small sample of schools for which ELA and mathematics proficiency rates for consecutive years are available, the foundation-supported schools appear to be making progress in the percentages of their students achieving state-mandated proficiency levels. In ELA, more than 55% of both new and redesigned schools had higher proficiency rate gains than their respective districts' average gains each academic year between 2002-03 and 2004-05. There were also gains in mathematics, although these unfortunately fluctuated widely in both types of schools from 1 year to the next.

Mathematics remains a particular area of concern. In future efforts to create new or redesigned schools, the foundation may wish to put additional emphasis on mathematics instruction and on strategies to improve student mathematics performance. The foundation is already supporting creative science, technology, and mathematics programs in specialized schools, and, if these schools are successful in improving students' mathematics performance, the redesigned schools could be encouraged to integrate lessons learned from these schools within their reform plans.

The foundation has also already commissioned a brief on *Improving Math Performance.* This brief discusses the experiences of several mathematics instructors in foundation-supported schools and points readers to useful tools and articles that might help teachers enhance students' performance in mathematics. The foundation may wish to disseminate this brief and other such publications to the grantees and the schools. The dissemination could be part of a broader effort to catalog and share a wide range of instructional support materials with the schools and districts.

Ultimately, to move students toward greater success, the foundation may wish to increase its focus on academic supports for students. As the early indicators suggest, the foundation has laid the groundwork for greater student success. Now the structural supports—new curricula, innovative instructional strategies, programs to get students who are behind their grade level up to their grade level—may warrant the same level of attention from the foundation.

Chapter 3. Making Instruction Rigorous and Relevant

Bill & Melinda Gates Foundation officials believe that to better serve all students, high schools need to become places that combine rigor in the academic program of every student (not just those in an honors or higher track) with relevance to their interests and potential career choices, supported by positive relationships that can inspire students both academically and personally. The foundation urges school leaders to make learning rigorous and relevant by doing the following:

- Building a culture of high expectations and academic challenge
- Aligning curriculum, instruction, and assessment with college admission standards
- Creating opportunities for in-depth exploration of topics
- Involving students in making decisions about what and how they learn
- Creating learning experiences that emphasize real-world connections and that relate to students' lives and aspirations
- Setting clear learning goals
- Providing intensive academic support so that all students perform to high levels
- Regularly assessing and providing informative feedback on student products
- Monitoring progress through multiple measures, including performance-based and standardized assessments

The relative emphases placed on these instructional strategies have shifted over time, with some gaining and others waning in importance in the foundation's grantmaking. However, the overriding focus on rigorous and relevant learning opportunities has remained constant.

This focus is important because research suggests that rigorous and relevant learning opportunities are likely to prompt intellectually complex student work (Bryk, Nagaoka, & Newmann, 2000; Newmann, Bryk, & Nagaoka, 2001; Mitchell et al., 2005). It suggests that when instruction makes high intellectual demands, students are more likely to rise to the occasion than when it does not and are more likely to produce work that reflects deep understanding, includes and explores new ideas, demonstrates reasoning and problem-solving skills, communicates well, and correctly applies language and mathematics conventions.[30]

The analyses that are reported in this chapter look behind the student work that was described in chapter 2 to test this theory, examining the rigor and relevance of classroom assignments and their relationship to the quality of students' work. Chapter 3 takes an in-depth look into the classrooms of foundation-supported new high schools, examining

[30] The research is guided by an analytic framework developed by researchers examining elementary and middle school reform begun under the Chicago Annenberg Challenge (Newmann, Lopez, & Bryk, 1998; Bryk et al., 2000; Newmann et al., 2001) and expanded to also address the foundation's tenets for high school teaching and learning (AIR/SRI, 2004; Mitchell et al., 2005).

instruction to see whether students in these schools have rigorous and relevant learning opportunities and comparing instruction in these schools with instruction in nearby comprehensive high schools without foundation ties.

Key Findings

The data in this chapter provide the basis for the following findings:

- Assignments in English/language arts (ELA) in the foundation-supported new high schools that we studied emphasized real-world settings and gave students a voice in shaping new learning opportunities. They also emphasized rigor. The ELA assignments that teachers gave in these schools were more relevant than classroom assignments in comprehensive high schools with similar student populations in the same or nearby districts, and they were equally rigorous.
- Many of the mathematics assignments that we gathered lacked rigor and relevance, whether they were from foundation-supported new high schools or other high schools. However, the mathematics assignments that teachers gave in the foundation's high schools were more rigorous and relevant than assignments from comparison high schools.
- When teachers assigned more rigorous and relevant assignments in ELA, students rose to the occasion and produced higher-quality work. In mathematics classrooms, the quality of student work was more strongly related to the rigor of assignments than to their relevance.
- Obstacles to effective mathematics instruction in the foundation-supported schools we visited included the need for high-quality curricula in mathematics, the need for teacher professional development and coaching in mathematics, weaknesses in the knowledge and skills that students brought with them to high school, and difficulty attracting certified and seasoned teachers in mathematics.
- Some foundation-supported schools have been working to address these obstacles by hiring mathematics coaches or working with university faculty as mentors and consultants on curriculum planning and instructional practice, sending teachers to workshops and conferences on mathematics instruction, and participating in national networks on mathematics improvement.

Table 5 summarizes the comparison of the relevance and rigor of instruction in foundation-supported new high schools and comprehensive high schools.

Compared with schools with similar student populations in the same or nearby districts, foundation-supported new high schools offered more relevant instruction in ELA and mathematics and provided equally rigorous instruction in ELA and more rigorous instruction in mathematics (Table 5).

Table 5. Instructional Outcomes in Foundation-Supported New High Schools, Compared With Comprehensive High Schools

	ELA	Mathematics
Relevance of assignments	+	+
Rigor of assignments	=	+

+: Statistically significant difference with new high schools scoring higher than comprehensive schools ($p < .05$)
=: No statistically significant difference between new and comprehensive high schools

This chapter describes these findings. The next section of this chapter describes the data that we collected and analyses that we conducted. The one that follows describes the structures that schools put in place to promote rigorous and relevant instruction. The third and fourth sections compare the rigor and relevance of instruction in foundation-supported new high schools with those in nearby high schools with similar student populations in ELA and mathematics, respectively. The fifth section examines the relationships between the rigor and relevance of instruction and the quality of students' classwork. The sixth section describes some of the factors that support or stand in the way of effective instruction in foundation-supported schools. The chapter concludes with a discussion of the implications that these data suggest.

Methods

Working with 10th-grade ELA and mathematics teachers in 12 foundation-supported new schools and 8 comparison schools with similar student populations in the same or nearby districts,[31] we collected two types of assignments: those that teachers considered typical of their day-to-day instruction and those that challenged their students to show what they knew and how they could perform at high levels. In schools using project-based learning methods, students' project proposals were sampled as assignments, along with teacher-developed assignments. For a subset of assignments and a random sample of students, we collected the student work associated with the assignments. The student work and the data that resulted from ratings of that work were described in chapter 2. We also interviewed and observed many of the teachers who provided assignments and work.[32] Additionally, we interviewed teachers in other foundation-supported new high schools and in redesigned high schools about teaching and learning.[33]

After the school year ended, we hired master teachers from other high schools to score assignments on the extent to which they provided students with rigorous, relevant learning opportunities. Each assignment was scored on the basis of the individual

[31] Data were collected in 2003-04 and 2004-05 for the 12 new schools and in 2002-03 for 1 comparison school and 2004-05 for 7 comparison schools.
[32] We interviewed 27 teachers from whom we collected assignments; in addition, we observed 24 of these teachers while they taught a class.
[33] In total, we interviewed 163 teachers in 35 schools.

components of the rigor and relevance scales, and many assignments were scored by multiple teachers. The scores on the different scales were combined to create one rigor score and one relevance score for each ELA and mathematics assignment. In each subject, the scores were converted to scales that described assignments as having little to no rigor, limited rigor, moderate rigor, or substantial rigor. The relevance scores were similarly converted. As described in chapter 2, the same procedures were used to evaluate the quality of student work. The percentages of assignments with little to no, limited, moderate, and substantial rigor in foundation-supported new high schools were compared with those in comparison high schools. The statistical significance of differences between the two groups of schools was examined by using an analysis method called hierarchical linear modeling (HLM). The models made adjustments for differences between schools in student and classroom characteristics. The same analysis methods and adjustments were used to examine the relevance data and to estimate the relationships between the characteristics of assignments and the quality of student work.

Data from the teacher interviews and observations that we conducted were classified according to the themes they addressed, and then the data were analyzed across all schools by theme. Table 6 provides an overview of data sources and comparisons made on the rigor and relevance of assignments.

Table 6. Data Sources and Analytic Approaches

Indicator	Data Source	Comparison
Assignment rigor	ELA and mathematics assignments from 12 new high schools and 8 comparison high schools	Between school types
Assignment relevance	ELA and mathematics assignments from 12 new high schools and 8 comparison high schools	Between school types
Aids and barriers to rigorous and relevant instruction	Interviews with 163 teachers in 35 high schools	Across schools

Additional details on the rigor and relevance scores, interview data, and the analyses on which this chapter relies appear in technical appendix C.

Promoting Rigor and Relevance

Foundation-sponsored schools have put a variety of structures in place to promote the rigor and relevance of their instruction. Some of these schools organize instruction around themes or individualized learning plans. Some support personalized instruction with structures such as advisories, multiyear teaching relationships, mentoring programs with business and community leaders, multidisciplinary courses, supplementary academic coaching, block scheduling, outside internships, and community service. School leaders have implemented these structures to help strengthen students' learning opportunities and to provide a basis for rigorous and relevant instruction. Table 7 lists criteria for evaluating the rigor and relevance of students' learning opportunities.

Table 7. Criteria for Instructional Rigor and Relevance

Rigor	Relevance
Extent to which assignments ask students to ♦ Move beyond the reproduction of information, taking what they already know and can do and using their knowledge and skills to create or explore new ideas ♦ Demonstrate conceptual understanding of important content ♦ Organize, interpret, evaluate, and synthesize information ♦ Communicate clearly and well	Extent to which assignments ask students to ♦ Address questions or problems with real-world applications ♦ Make choices about what they will study and how they will study it ♦ Take on plausible writing roles and submit their work to real audiences

Exhibit 1 presents a mathematics assignment that meets many of these criteria for substantial rigor and relevance. The exhibit gives directions to the teacher for organizing and then administering the assignment to his or her students.

Exhibit 1. Mathematics Assignment With Substantial Rigor and Relevance

Paying for Your Dream

Framing Questions

- How much does charging really cost me?
- How can Excel help me see different alternatives?

Reason

Pose the question: Could you afford anything on your billion-dollar spending spree? What if you had a credit card?

Risk

Vote on how long it would take you to buy a certain item using a credit card and how much interest it would cost you.

Choose someone's expense in the room. Say, for example, a student wanted to buy a $3,000 plasma screen TV. Most credit card companies will let you pay as little as 2% of the price each month, so that's (2/100) x ($3,000) = $60/month. How long do you think it would take to pay it back: less than a year? Two or three years? Five years? More than ten years? Have students vote.

Rethink

Set up a spreadsheet on the computer to calculate the credit card debt. Then try different scenarios to find ways to pay off the debt faster.

Students should change the interest rate and payment schedule to see how small changes can lead to big effects later. For example, how would one payment of $500 help if you did it right at the beginning? How would it change if you paid $100/month? What if you paid $50/month? What if the interest rate was 12% instead of 18%?

Reflect

What surprised you about the credit card pay-off plans? How does this activity change (or why does it not change) your view about credit cards? Start by having students write about the different ways students could pay off their loans. The following two questions can help guide their writing.

How does changing the interest rate change how long it takes to pay off the loan? How does it change how much interest you pay during the life of the loan?

How does changing the payment schedule change how long it takes to pay off the loan? How does it change how much interest you pay during the life of the loan?

This assignment is considered rigorous because it calls for students to demonstrate conceptual understanding of variables, expressions, and equations—central ideas in Algebra. It asks students to model the compounding-interest problem and to manipulate variables within the model to see the effects of different interest rates and payment schedules. The assignment calls for problem solving and reasoning in a problem setting that is fairly complex for a 10th grader. It requires students to generalize and make predictions about the nature of compounding interest and to reflect on the results of the modeling activity. Though students make few choices in the course of their work on this assignment, it scores well on relevance because the mathematics content has a real-life application.

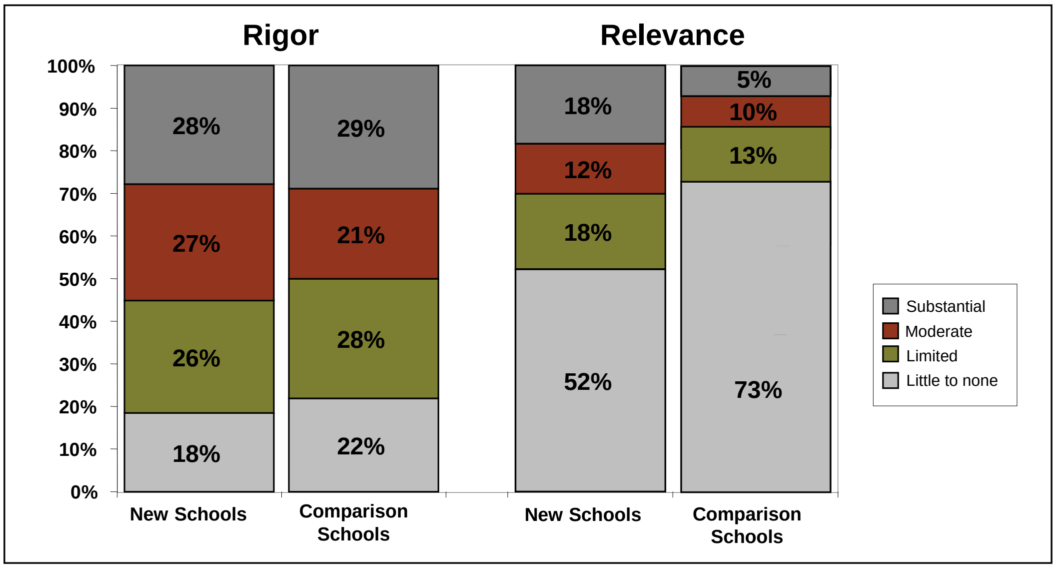
Discussion of the Findings

Assignments in ELA in the foundation-supported new high schools that we studied emphasized real-world settings and gave students a voice in shaping new learning opportunities. They also emphasized rigor. The ELA assignments that teachers gave in these schools were more relevant than classroom assignments in comprehensive high schools with similar student populations in the same or nearby districts, and they were equally rigorous.

Figure 9 compares the rigor and relevance of the ELA assignments that were given in foundation-sponsored new high schools with the rigor and relevance of assignments given in nearby high schools with similar student populations and without foundation ties. The graphs in Figure 9 show the extent to which 10th-grade students in new high schools and local comparison schools received rigorous assignments that were intellectually demanding and relevant assignments that prompted them to make real-world connections and choices about their work.

ELA assignments in foundation-sponsored new high schools showed higher levels of relevance than assignments in nearby comparison schools. Levels of rigor were similar in the two types of schools (Figure 9).

Figure 9. Rigor and Relevance of 10th-Grade ELA Assignments in Foundation-Sponsored New High Schools and Nearby Comparison Schools



Data sources. Assignments from teachers in 12 new schools (331 assignments from 33 teachers) and 8 comparison schools (127 assignments from 19 teachers)

Figure 9 shows that the 10th-grade ELA assignments given by teachers in new high schools were as rigorous as those given by faculty in nearby comparison schools, and more relevant. In both sets of schools, about half of the ELA assignments were judged to have either moderate or substantial rigor. The right-hand graph shows that 30% of the ELA assignments submitted by faculty in new high schools were judged as having moderate or substantial relevance, compared with only 15% of assignments from faculty in comparison schools. The difference between the relevance of assignments given in new and comparison schools is statistically significant.[34,35]

When 2003-04 ELA rigor scores for the 12 new schools were compared with 2004-05 scores for the same schools, there was no significant difference. The same was true for ELA relevance scores.

Exhibit 2 provides an example of an ELA assignment with substantial rigor and relevance.

[34] HLM analyses controlled for differences in school demographics and students' eighth- or ninth-grade test scores.

[35] One should exercise caution in attributing differences in the rigor and relevance of learning opportunities between new and comparison high schools entirely to the reform efforts. It is possible that, in addition to the reform efforts, schools differed systematically in other ways that could have influenced instructional practices. Even with adjustments for differences between schools in student and classroom characteristics, there could well be unmeasured differences that are related to differences in instructional practice across school types.

Exhibit 2. ELA Assignment With Substantial Rigor and Relevance

***MACBETH* ACT II, Writing Topics**

Use one of the following prompts to write about the events and characters of Act II of *Macbeth*.

- You are a director who is staging or filming Macbeth. Cast the part of Macbeth with a modern actor, and write a note to this actor telling him why Macbeth murdered Duncan. Tell the actor how you want him to play the part.
- You are a defense lawyer representing Duncan's sons, Malcolm and Donalbain, who are on trial for the murder of their father. Write your opening statement to the jury, explaining why Malcolm and Donalbain are innocent.
- You are a staff writer working for Lady Macbeth, who is the publisher of a woman's magazine like Martha Stewart's Living. Write a "How-to" piece for the first issue of Lady Macbeth's magazine explaining how a woman can motivate her husband to commit a murder.
- You are a detective piecing together the real truth about Duncan's murder. You have the murder weapon, Duncan's body, and the bodies of the servants. Using other evidence that was subsequently found (Macbeth's letter to Lady Macbeth, Banquo's statements about the witches' prophecies), build a case for Macbeth's guilt in the murder of Duncan.

Next, we repeat these analyses with mathematics assignments. The story is different in mathematics and in ELA.

Many of the mathematics assignments that we gathered lacked rigor and relevance, whether they were from foundation-supported new high schools or other high schools. However, the mathematics assignments that teachers gave in the foundation's high schools were more rigorous and relevant than assignments from comparison high schools.

Figure 10 presents data on the rigor and relevance of mathematics assignments in foundation-sponsored new high schools and other large high schools.[36] The data show that only small percentages of mathematics assignments in the foundation's new high schools and in other high schools were rigorous. Fewer than 20% of the mathematics assignments we collected from new and comparison high schools had moderate or substantial rigor. However, 42% of the mathematics assignments submitted by teachers in new high schools were judged to have at least limited rigor, compared with only 31% of mathematics assignments from comparison schools. This difference between the two school types is statistically significant.[37]

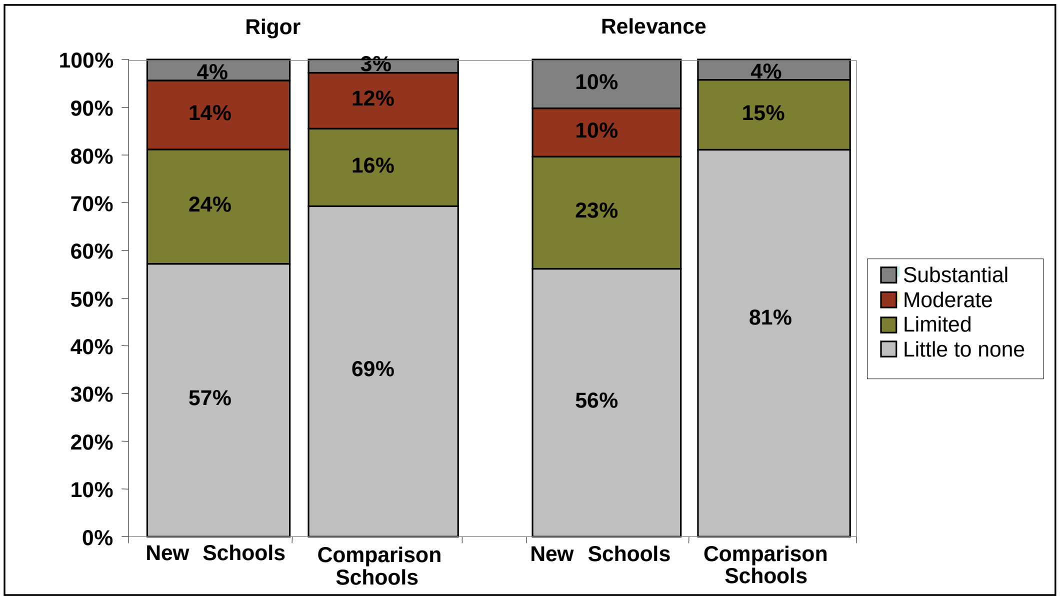
---

[36] The rigor of mathematics assignments was estimated independently of the rigor of ELA assignments, and the scales are not comparable. The same is true for the relevance scales.

[37] In *Rigor, Relevance, and Results: The Quality of Teacher Assignments and Student Work in New and Conventional High Schools* (Mitchell et al., 2005), we described an analysis that found that mathematics assignments in the new schools were less rigorous than those in a set of 16 large high schools in the state of Washington. Although this year's assignments came from the same set of new schools as was used in last year's analysis, the comparison schools are a new school sample, which consists of the schools that the students in the new schools would have attended if the new schools had not been formed. Also, a new set of assignments was collected from the new schools and combined with the assignments from the previous year. Although we have not made

Mathematics assignments in foundation-sponsored new high schools showed higher levels of rigor and relevance than assignments in nearby comparison schools (Figure 10).

Figure 10. Rigor and Relevance of 10th-Grade Mathematics Assignments in Foundation-Sponsored New High Schools and Nearby Comparison Schools



Data sources. Assignments from teachers in 12 new schools (232 assignments from 33 teachers) and 8 comparison schools (135 assignments from 19 teachers)

The same pattern holds for the relevance of mathematics assignments. Only small percentages of mathematics assignments were judged to be relevant, but mathematics assignments in foundation-sponsored new high schools were significantly more relevant than those in large high schools nearby. Twenty percent of the mathematics assignments given by faculty in new high schools were judged to have moderate or substantial relevance, compared with only 4% of mathematics assignments from comparison schools.

When 2003-04 mathematics rigor scores for the 12 new schools were compared with 2004-05 scores, there was no significant difference. The math relevance scores in schools decreased, however, from 2003-04 to 2004-05.

Exhibit 3 presents a mathematics assignment with limited rigor and limited relevance. The assignment focuses on an important mathematical idea—using data to create direct and inverse variation models in equations and graphs, and then using these models to make predictions. Examined with our rubrics, this assignment has several shortcomings. It scores relatively low on rigor because it primarily requires students to demonstrate

---

the comparison directly, the two analyses suggest that mathematics assignment rigor is lower in this new sample of comparison schools than in the Washington high schools. This year's analysis uses a more appropriate comparison group.

procedural knowledge. The problems are fairly basic and routine for 10th graders, require little problem solving, and could be completed by applying previously learned procedures. This assignment requires little more than numerical answers or graphs, and could be made more rigorous by asking for supporting work, explanation, or justification. For example, problem 2 could ask how the three equations are similar, how they are different, and why. The assignment does attempt to relate the mathematics to the real world, but the work is not done for an authentic purpose, so it does not receive the highest relevance score. There is no student involvement in developing the assignment; the topic and the problems are prescribed.

Exhibit 3. Mathematics Assignment With Limited Rigor and Relevance

Unit 2 Preview

1. The height of a ball thrown up in the air from atop a wall 200 feet high is given by the equation: $h = -16t^2 + 96t + 200$

   a. Find a reasonable domain and range for the function.
   b. Find an approximate viewing window for the function and graph it.
   c. What is the highest point the ball reaches?
   d. At what time does the ball reach the ground?

2. Graph the following equations:

   a. $y = x + 3$
   b. $y = (-1/3)x - 2$
   c. $2x + 3y = 9$

3. If the points (3, 5) and (-2, 1) are two points on a line, find the following information:

   a. The slope of the line that passes through the two points.
   b. The vertical intercept of the line in part a.
   c. Graph the line; include at least four labeled points.

4. The following data table represents the measurements taken from the relationship between the number of Big Macs eaten and calories consumed.

Number of Big Macs	Calories
5	1,500
7	2,100
3	900
9	2,700

   a. Determine which mathematical model fits the data.
   b. Write an equation.
   c. Determine how many calories will be consumed if you eat 12 Big Macs.
   d. How many Big Macs were eaten if 1800 calories were consumed?

Exhibit 3. Mathematics Assignment With Limited Rigor and Relevance *(continued)*

5. The following data table represents the measurements taken showing the relationship between the weight of a person and the distance from the center of a seesaw that person is sitting to have it balance.

Weight (pounds)	Distance (feet)
100	5.5
150	3.67
200	2.75
220	2.5

   a. Determine which mathematical model fits the data.
   b. Write an equation.
   c. Determine the distance from the center if a person weighs 180 pounds.
   d. How much does a person weigh if they have to sit 4 feet from the center?

6. The amount of water in a spherical globe varies directly with the cube of the radius. When the radius is 3 cm, the volume of water is 113.04 cubic centimeters.

   a. Determine an appropriate equation for this situation.
   b. Determine the amount of water needed for a globe with a 12 cm radius.
   c. How large is the radius if the globe requires 14,130 cubic centimeters of water?
   d. Graph the equation showing the basic shape of the equation.

When teachers assigned more rigorous and relevant assignments in ELA, students rose to the occasion and produced higher-quality work. In mathematics classrooms, the quality of student work was more strongly related to the rigor of assignments than to their relevance.

Figures 9 and 10 provided data on the characteristics of ELA and mathematics assignments in new and comparison high schools. As noted earlier, prior research (Bryk et al., 2000; Mitchell et al., 2005) suggests that rigorous and relevant learning opportunities are more likely to prompt high-quality student work than are assignments that make fewer demands and offer students fewer real-world connections. Here, we reexamine this relationship and document the relative importance of rigor and relevance to the ultimate quality of students' work.

When we examine the relationships between the characteristics of teacher assignments that were just described and the quality of student work that was reported in chapter 2, we find that rigor and relevance have differential effects on the quality of students' work in ELA and mathematics.[38] In ELA, both more rigorous assignments and more relevant

[38] In both subjects, rigor and relevance are individually positively correlated with student work quality. But when we use analysis methods that allow us to simultaneously examine the relationships between rigor and relevance and student work in mathematics, we see that the relationship between relevance and student work quality is almost entirely explained by the positive relationship between relevance and rigor. That is, the relationship between relevance and student work quality in mathematics operates through the strong positive relationship between rigor and student work quality in mathematics. The path coefficients between rigor and student work quality and relevance and student work quality in mathematics are .79 and -.02. The relationships are different

assignments are more likely to prompt high-quality student work. In mathematics, however, it is rigor that makes the difference in prompting high-quality student work. Mathematics assignments that are more rigorous are more likely to prompt high-quality student work than mathematics assignments that are not. Though rigor and relevance often go hand in hand in mathematics assignments, according to our data, without rigor, real-world tasks are no more likely to be associated with high-quality student work than textbook treatments of the same content. See technical appendix C for additional detail on these findings.

When we combine these findings about the relationships between the characteristics of assignments and student work with the data we reported in chapter 2 about the differences between the quality of students' work in foundation-supported new high schools and comparison schools, the story is clear for ELA, but interpretive difficulties arise for mathematics. ELA assignments were equally rigorous in new and comparison high schools but significantly more relevant in the new high schools. Both rigor and relevance had positive effects on the quality of the work students produced. Students at foundation-supported new schools produced higher-quality work in ELA than students in comparison schools did.

In mathematics, the story is more complex. Although very few of the mathematics assignments given in foundation-supported new high schools or comparison schools had substantial rigor, more of the mathematics assignments submitted by the new high schools than by comparison high schools had at least limited rigor. Mathematics assignments from new high schools also were more relevant than assignments from comparison schools. According to our analyses, substantially and moderately rigorous mathematics assignments are more likely than assignments with limited rigor to prompt high-quality student work. But, as we reported in chapter 2, there were no significant differences in the quality of students' work in mathematics between foundation-supported new high schools and comparison schools, and, as already mentioned, few mathematics assignments had moderate or substantial rigor. This conundrum and the limited rigor of mathematics assignments in foundation-supported high schools are discussed in the context of the obstacles described next.

Obstacles to effective mathematics instruction in the foundation-supported schools we visited included the need for high-quality curricula in mathematics, the need for teacher professional development and coaching in mathematics, weaknesses in the knowledge and skills that students brought with them to high school, and difficulty attracting certified and seasoned teachers in mathematics.

To better understand what supports and inhibits rigorous mathematics instruction in reforming high schools, we examined data from teacher interviews in foundation-supported new and redesigned high schools.[39] Data on instructional resources for

in ELA. Rigor and relevance each have independent effects on student work quality; the path coefficients for rigor and student work quality and relevance and student work quality in ELA are .64 and .54, respectively. These analyses included data for both new and comparison high schools.

[39] We interviewed 163 teachers, a substantial number of whom talked about mathematics instruction.

mathematics, teachers' instructional strategies in mathematics, and classroom challenges were examined. Analyses focused on the conditions and resources that inhibit rigorous and relevant instruction in mathematics.

Our interviews and observations in foundation-supported schools revealed that faculty used a wide range of resources in their mathematics programs. Some of these resources were locally developed or provided by grantees; others were Web accessible or commercially available. In some schools, faculty used a combination of locally developed and commercial or other externally provided resources to make up their mathematics curriculum. In the sample of foundation-supported schools that were visited in 2003-04 and 2004-05, the more widely available mathematics resources used by faculty were Accelerated Math, Boxer Math, materials from the Buck Institute, Cognitive Tutor, College Preparatory Mathematics, Discovering Geometry, Geometer's Sketchpad, the Interactive Mathematics Program, and products from Riverdeep.

Our conversations with faculty about these resources revealed that teachers were not uniformly satisfied with them. Teachers said that they had difficulty engaging their students with some of these materials, particularly the more procedurally oriented ones. They described some materials as text heavy and difficult to use with students with poor reading skills. Many faculty also lamented the shortage of resources for project-based instruction in mathematics. (Exhibit 4 shows a proposal for a student mathematics project that was locally developed.) Faculty concluded by saying that they had little support for the use of these resources or for local development of instructional materials in mathematics.

Exhibit 4. Student-Generated Project Proposal in Mathematics

This is a student-generated proposal for a project that is related to the student's internship experience. The proposal is the product of negotiations between the student, his faculty advisor, and his internship mentor. It describes the instructional objectives for the project and the way that the student will demonstrate mastery of the learning goals that the student, advisor, and mentor have agreed on. It argues that the project and the resulting product are likely to meet the school's standards for rigor and relevance.

3–D Model

Introduction

The project that I will be working in this quarter is a 3 dimensional model of the [name] School addition that the [name] Company is constructing. I will show an example of how the school is going to look like once it's completed. What I'll be showing at my exhibition is my model, overview of the whole construction, and live photos of the whole building or some parts of it, and maybe I'll give a tour. My project will benefit my [internship] site because they have blueprints and not a 3 dimensional model.

Focus

What I hope to learn from this internship/project is more about the construction process. My research paper will be about the process of building. What follows are my learning goals. Quantitative Reasoning fits in because there are a lot of measurements that you have to make sure are correct because if you don't the building, in my case, would either fall about apart because there would be no balance or the pieces wouldn't really go that well together. Also if

Exhibit 4. Student-Generated Project Proposal in Mathematics *(continued)*

for an example that has happened to me couple of times is that if you don't get the perfect measurements then one side of the building will be shorter or taller and the building model couldn't be glued together properly. Empirical Reasoning fits in as well because there is a part of the building that doesn't exist and it's my job to come up with a reasonable elevation. Communication is another goal because I have to communicate with my mentor every time I'm there to know what I have to do and if I have done things the correct way. I also translate the blueprints to people that don't [know] how to read them. For Social Reasoning I have been finding out about the history of the [name] School construction. Personal Qualities fits in as well as helped me because you have to have a lot of patience in order to get all the blueprints together and having to find one you want which isn't easy if you don't know how to read them. Persistence is another personal quality because if there is a lot of posting [putting the building parts in their appropriate order] to do you have to keep up with what you did the last time that you were there. Those are the main learning goals that I will be using for this quarter's project.

Authenticity

I chose to participate in this internship because it's seriously what I'm interested in (it might be a possible career) next to all the other things I am interested in. I'm interested in construction the most because I have learned about it from my father who is a construction worker. When he doesn't have his partners that work with him, I go tagging along with him. While I'm doing whatever it is that he needs help with, I enjoy it very much, and the thing about it is that neither he nor I get bored doing it.

The way that this project will expand my learning is because it will teach me something different that I could and most likely [would] use in the future as well as expanding my skills. I can say that my project is challenging because you have to be able to read blueprints and scale which I could proudly say I know how to do. The resources that I will use for this project are drawings, professional staff and the field work, as I will be seeing how the actual building is coming out. What I will be creating as a model for this project, is a 3 dimensional model of the [name] School and the elevations. My project is real because it's a model of an actual building being built. The people that will benefit from this project will be staff and [name] Mentor Program. I'll know that my model is completed once I compare it to the blueprints and I'll ask [name] to see if it's met high standards.

The final due date will be the 27th of May. The steps that I will be taking in order to complete this project are, find a project that will show my skills and creativity; find a blueprint that will give an idea of what to build; color the blueprints so that we know all the different types of materials we need in order to get the model together; buy the materials we needed; make all the measurements, cut out the walls out of the poster board, put it all together and start adding the detail.

I will document my hands-on learning by showing my model, explaining what I've done, and describing my research in a paper.

Some teachers said that they needed help in developing the expertise and strategies that are necessary to teach mathematics effectively. Faculty said that they wanted and needed additional professional development and ongoing coaching in mathematics. In particular, teachers wanted help in learning how to provide students with sufficient scaffolding to develop conceptual understanding of mathematics content, to solve problems and reason in mathematics, and to communicate effectively about mathematics.

With respect to instructional supports, many teachers also lamented the lack of time to take advantage of the teacher professional development offerings that were available or to make the changes suggested by the professional development they received. One teacher explained, "We've had a lot of staff development, but it doesn't really help us to know what we have to do and not have the time to do it. And we're all kind of overwhelmed with that."

The teachers with whom we spoke described clear weaknesses in students' mathematics backgrounds as another obstacle to rigorous mathematics instruction in foundation-supported schools. Teachers reported (and the earlier-reported achievement test data confirm) that many of the students who walk in the doors of foundation-supported schools come with weak mathematics backgrounds. Many teachers described the difficulties of teaching Algebra II and trigonometry to students whose mathematics skills are limited to basic arithmetic. One teacher said, "It's kind of hard to understand the frustration of teaching a student whose skill level is basic math, to teach them logarithms!"

In addition, many of these schools use heterogeneous grouping, which means that students with a wide range of knowledge and skills are taught in the same classrooms. This practice is a particular challenge in mathematics because mathematics has a more tightly prescribed scope and sequence than many other subjects. Procedures and concepts build on each other to a much greater degree than is the case, for example, in ELA. Few faculty have training in teaching heterogeneous classes in mathematics.

Undoubtedly, the limited rigor and relevance of mathematics assignments in foundation-supported schools contribute to students' low-quality work in mathematics. But even when faculty give assignments that are rigorous and relevant, the lack of support they receive, students' weak mathematics skills when they begin high school, and the wide range of skills present in many teachers' classrooms confound teachers' efforts to help students work to high levels. In the absence of better mathematics curricula and proper professional support for teaching mathematics to underprepared students, lackluster student work is likely to result.

This challenging scenario in mathematics is exacerbated by a problem that troubles many high-poverty schools across the country. Several of the new high schools that we studied had trouble attracting mathematics faculty to their schools. The teachers with whom we spoke said that attracting experienced mathematics teachers is particularly difficult, and retention rates for mathematics faculty are uncomfortably low. As a result, some of these schools have an ongoing need to train new mathematics teachers. One faculty member said:

> [Our] new math teachers this year [are] great people in terms of their passion and desire. But they needed training, and they definitely needed to understand [our school model]. They needed to understand how important it is to teach to the standards, and they also needed more support in terms of how to teach Algebra II to students who aren't really good at Algebra I. Our students

> sometimes don't have the basics. You need really seasoned teachers with a lot of passion and a lot of wisdom and a lot of strength in those areas to push kids. [The new teachers] really struggled in terms of getting our kids to where they needed to be.

Tables 8 and 9 summarize obstacles teachers reported and strategies they have used or would like to use to improve mathematics instruction and learning.

Table 8. Obstacles to Effective Mathematics Instruction, Reported by Teachers

- Need for high-quality curricula
- Need for professional development and coaching
- Need for coaching on scaffolding for students with weak mathematics backgrounds
- Need for instruction on teaching mathematics to heterogeneous classes
- Difficulty attracting and retaining certified and experienced teachers
- Shortage of resources for integrating mathematics into project-based instruction
- Need for more support in development of instructional materials
- Lack of time to take advantage of professional development

Table 9. Strategies for Improving Mathematics Instruction, Reported by Teachers

- Attend conferences and workshops about best practices for instructional support and content in mathematics
- Participate in teacher networks about mathematics instruction
- Hire coaches or consultants to provide school-based workshops and curriculum supports
- Partner with university staff for mathematics support
- Appoint lead mathematics instructors to mentor other faculty
- Encourage interdisciplinary collaboration and project planning
- Institute common planning time
- Give students a double dose of mathematics instruction
- Supplement project-based instruction with more systematic treatments of content

Some foundation-supported schools have been working to address these obstacles by hiring mathematics coaches or working with university faculty as mentors and consultants on curriculum planning and instructional practice, sending teachers to workshops and conferences on mathematics instruction, and participating in national networks on mathematics improvement.

Some foundation-supported schools and the grantee organizations that work with them are trying to meet these mathematics challenges head on. In some cases, the grantee organizations have created support systems for mathematics teachers. There are grantees, for example, that host annual conferences to share best practices for teachers from their sponsored schools. These conferences often have sessions designed around mathematics content and various instructional supports for in-service teachers. Other grantees provide opportunities for teachers from different schools to communicate through mini-conferences or workshops.

Additionally, some grantees have hired mathematics coaches for their affiliated schools. These coaches generally have several years of mathematics teaching experience and provide direct professional development to teachers in the form of weekly or monthly workshops, curriculum support, observations, or departmental meetings. Interviewed teachers referred to grantee support as helpful for their instruction, although the ultimate impact on student achievement has yet to be determined.

Faculty at a few foundation-supported schools themselves have used funds from grantees' professional development accounts or grant money from private sources to hire outside consultants to provide on-site support. The consultants' routines vary, with some serving full-time at a school for an entire year while others come to schools intermittently, holding workshops. There are also schools that have partnered with local universities to hire professors to work with their staff as a whole or with individual math teachers. Although the particular mechanisms for providing on-site support vary from school to school, the shared goal across campuses is to provide teachers with on-site development that fits their particular needs.

A small number of schools have designated lead mathematics teachers who serve as mentors and supporters for other mathematics teachers in the school. These lead teachers host meetings on instructional topics, observe teachers, or conduct demonstration lessons in teachers' classrooms.

Additionally, some school leaders have engaged the entire school staff in mathematics professional development workshops or curriculum training to encourage interdisciplinary collaboration and project design. Furthermore, conscious of the need to give teachers time to improve their curricula and lesson plans, some schools' mathematics departments have instituted common planning time for teachers. During this common planning time, mathematics teachers come together to ensure that they are integrating concepts across instructional years or that they are interpreting curricula in similar ways. The teachers with whom we spoke noted that common planning time was a key enabler to helping them teach effectively.

A few of the schools have doubled the time in their instructional days that students spend doing mathematics. In some schools that follow project-based learning models, faculty are supplementing students' project-based mathematics work with software-based mathematics programs like those described earlier. One teacher explained, "I try to do skill-based things 2-3 days a week because of their weak foundation skills. The school administration wants to do projects, but without the skills it's hard to do." Many of the teachers and students who use these types of software packages express dissatisfaction, saying that the problems make few real-world connections. Despite these limitations, faculty say that it is important to supplement students' project-based work in mathematics with more systematic treatments of the content, even in schools that use little direct instruction.

Implications for the Initiative

The data that we collected from ELA classrooms suggest that foundation-supported new high schools and their students are having relatively good success in comparison with large comprehensive high schools and their students. The assignments that teachers gave in new high schools were more relevant than and as rigorous as assignments given in neighboring high schools without foundation ties. The assignments prompted higher-quality work from students in foundation-supported high schools.

These results suggest that a more careful analysis of the curricular emphases and instructional practices that are used in the more successful high schools might provide lessons for instructional improvement in schools that are struggling.

The implications of the mathematics data are clear. Foundation-supported schools are facing the same problems incorporating rigor in mathematics as other high-poverty schools across the country. The results suggest that a search is needed for mathematics practices that hold promise for improving mathematics learning in high-poverty high schools. More ambitiously, the foundation should consider investing in the identification and optimal sequencing of the mathematics knowledge and skills that are the most critical for students who are achieving below grade level to master. The identification and design of instructional activities that make critical knowledge and skills accessible to underprepared students would be a valuable contribution.

Identifying or developing resources that make these same mathematics concepts accessible to faculty is also likely to be helpful. Teachers report that they need descriptions, modeling, and coaching to improve their practices in teaching mathematics. They also need support for building the scaffolding necessary to improve the achievements of students who have weak mathematics backgrounds.

Chapter 4. Partnering With Districts to Offer Portfolios of High School Options

The introduction to this report described the Bill & Melinda Gates Foundation's call for secondary schools with rigorous academic programs that graduate students ready for college, family-wage jobs, and the demands of good citizenship. It discussed the foundation's challenge to underperforming school systems across the country to increase their college-ready graduation rates for low-income, African-American, and Hispanic/Latino students to 80% by 2014 (Vander Ark, 2003). For many districts where poor and racial/ethnic minority students are in the majority, the foundation's challenge effectively requires schools to double current high school graduation rates and quadruple the numbers of graduates ready for postsecondary education.

Though the foundation's prescriptions for meeting this challenge have evolved over time, their central element has been the conviction that students and families need choices among schools that support students' differing needs and interests (Vander Ark, 2003; Hill, 2006; Bill & Melinda Gates Foundation, 2004; Bill & Melinda Gates Foundation, 2005a). Foundation officials argue that districts should build portfolios of schooling options so that families can select the schools that best meet students' needs. Foundation leaders contend that school portfolios should include a range of schooling options, including choices among schools that teach traditional subjects; schools that organize curricula around particular themes, such as science and technology or social justice; and schools where students have individualized learning plans and work on carefully designed projects (Bill & Melinda Gates Foundation, n.d.-b).

Like reformers before them, foundation officials believe that supportive operating environments for school choice are unlikely to develop in districts where schooling options are few (Mirel, 2001; Parthenon Group, 2002). They contend that until 30% of a community's schools offer enrollment choices to students and families, a district is unlikely to institutionalize the policies and practices that are needed to support varied schooling options. Consequently, the foundation urges policymakers in the districts where it works to provide school choice in at least 30% of their schools.

In the districts where it works, the foundation seeks policies that challenge schools to become high-performing organizations and encourages practices that support schools as they build their capacity to reach this goal (Bill & Melinda Gates Foundation, 2005a). Foundation officials argue that districts create high-challenge environments when they commit to college-ready missions, effective governance, strong accountability, and engaged communities. The foundation asks districts to foster change by creating support systems focused on curricula, teachers, schools, and students and families. Exhibit 5 describes the challenges and supports that the foundation looks for in the districts where it helps to create and redesign high schools (Bill & Melinda Gates Foundation, 2005a).

Exhibit 5. Characteristics of High-Performing School Districts

High Challenge	
Organizational mission	Common standards that prepare all students for postsecondary education, work, and citizenship
Governance	Stable, effective local governance focused on results and equity empowering improvement with transparency, measurement, and responsiveness, as well as alignment with state goals and policies
Accountability framework	Transparent performance management system with steps of progressive intervention that provides support for all—students, staff, school, and system—and is relevant to the challenge
Community support	Proactive strategies to engage parents, citizens, and business and civic leaders—resulting in an informed community that makes quality education a priority
High Support	
Curriculum support	Learning expectations that provide a spine for instructional materials, diagnostic assessments, ramp-up supports for students, and teacher development activities
School support	Strong learning and support networks for schools; dollars that follow students and reflect student needs, creating budgets that allow school-based decision-making; effective core services provided; optional purchased services available
Teacher support	Instructional leaders identified and developed; district recruiting; school/network-based hiring; 3-year induction with ongoing, job-embedded development in a professional learning community; compensation that reflects ability, performance, and responsibility
Student support	An advocate for every student who ensures appropriate guidance and academic support and connection to family services; frequent and specific performance feedback against clear expectations for every student

This chapter examines the foundation's work in school districts across the country, comparing reformers' efforts with the foundation's targeted goals. Although we lack evidence on graduation rates in the foundation's sites (only a handful of foundation-supported schools have been in operation long enough to have graduating classes), we examine the availability of diverse schooling options in districts supported by the foundation. We also examine the policies and practices that are in place to support these schools. The chapter describes the foundation's aspirations for high-performing districts and gives detail on the assistance that districts provide to reforming schools. It also describes barriers to reformers' work and discusses the future prospects of some of the schools and instructional models that the foundation has helped to create. The chapter ends with a discussion of the implications of these results for the foundation's future work.

Key Findings

The findings in this chapter are based on data from 17 districts where the foundation is working with high schools. The data suggest the following:

- By 2004-05, 15 of the 17 districts included in our evaluation provided school choice to students and families in 30% or more of their high schools. Across the 17 districts, a third of the high school students attended schools of choice.
- Students and families in the 17 districts could choose from among several small, focused high schools in large high school buildings or new free-standing high schools that represent different educational philosophies, curricular emphases, and instructional models. Across the 17 districts, 28% of the high school students attended schools that took theme-based or other innovative approaches to curriculum and instruction.
- In some of the school systems that we studied, district and school staff developed their own school models, designing curriculum and instruction for their schools and setting up the academic structures for them. In others, district and school staff adopted school models that were supported by school franchises and other external providers. Many districts put in place a mix of locally developed and externally provided school models.
- There was evidence of some of the challenges and supports the foundation recommends in all of the districts that we studied. These included college-ready missions; effective governance; strong accountability; community engagement; and supports for curricula, teachers, schools, and students and families. None of the districts had all the recommended challenges and supports in place.
- Some schools in these districts had uncertain futures. By 2004-05, faculty at 8 of the 12 third- and fourth-year new schools in our evaluation felt confident that their schools were here to stay. Additionally, faculty at three of the five third- and fourth-year redesigned schools we visited predicted that the instructional models that defined their small learning communities would remain intact in 2005-06. In the other new and redesigned schools, some faculty members were uncertain about the futures of their programs.
- Whereas teacher turnover in urban high schools across the country averages 18% per year, 30% or more of the teaching staff in the foundation-supported new schools left after each of the schools' first 3 years of operation.

Table 10 provides a snapshot of system outcomes where the foundation has helped to create and redesign high schools.

By 2004-05, almost all of the districts included in our evaluation provided school choice to students and families in 30% or more of the high schools. All of the districts provided choice among schools with different educational models (Table 10).

Table 10. Snapshot of System Outcomes in 17 Districts Where the Foundation Has Helped to Create and Redesign High Schools

Outcome	Number of Districts
Met the foundation's target for 30% schools of choice	15 of 17 districts
Provided choice among schools with different educational models	17 of 17 districts

This chapter details each of these findings, describes the strategies that districts use to promote and manage change, and discusses the future of third- and fourth-year schools in selected districts.

Methods

This chapter focuses on 17 districts that were in the third or fourth year of their work with the foundation by 2004-05 and had one or more schools included in our evaluation; most of these districts had multiple foundation-supported schools in operation in 2004-05.[40] These districts are located in 11 states.[41] This chapter gives particular attention to the schools in these districts that were in their third or fourth year of operation or redesign in 2004-05 and for which we have 3 or 4 years of evaluation data. Included in this group are 12 new high schools and 5 redesigned high schools located in 13 of the 17 districts.

The data for this chapter came from interviews with district officials, staff at foundation-supported schools, and staff from the reform organizations that support the schools' work. Additional information came from district and school records and from district and school Web sites. Table 11 lists the system and school outcomes that are described in this chapter and their corresponding data.

[40] The 17 districts are Anderson Union, Baltimore, Chicago, Cincinnati, Cotati-Rohnert Park, Denver, El Dorado, Milwaukee, New York City (Bronx, Brooklyn, and Manhattan), Oakland, Oklahoma City, Providence, Ravenswood, Sacramento, St. Paul, West Clermont, and Worcester.
[41] Foundation investments in the 17 districts ranged from less than $1 million to more than $115 million (Foundation Strategy Group, 2005). Investment levels in the 17 districts spanned a wide range, in part because the districts' enrollment levels vary so much. The smallest district in this group enrolled approximately 2,300 high school students in 2004-05; the largest enrolled more than 300,000 students. Schools in 6 of the 17 districts had grants that were less than $1 million; 6 districts had grants that ranged from $1 million to $10 million; and 5 districts had grants that were more than $10 million.

Outcome data for school systems were obtained from district records and interviews with reform organization staff, district officials, and school staff (Table 11).

Table 11. Outcomes and Data Sources

Outcome	Data Source
Numbers of districts with at least 30% schools of choice	Records from 17 districts
Numbers of schools following different educational philosophies, curricular emphases, and instructional models	Records from 17 districts
Numbers of school choices available to students	Records from 17 districts
District supports for high school portfolios	A total of 309 interviews with district officials (42), reform organization staff (66), teachers (127), and other school leaders (74) in the 17 districts
Predicted sustainability of the foundation-supported schools and instructional models	A total of 126 interviews with teachers (75) and other school leaders (51) at 12 new schools and 5 redesigned schools in 13 of the 17 districts

Additional detail about the district and school data sets and the analyses on which this chapter relies appears in the technical appendix D.

Discussion of the Findings

Districts and School Choice

By 2004-05, 15 of the 17 districts included in our evaluation provided school choice to students and families in 30% or more of their high schools. Across the 17 districts, a third of the high school students attended schools of choice.

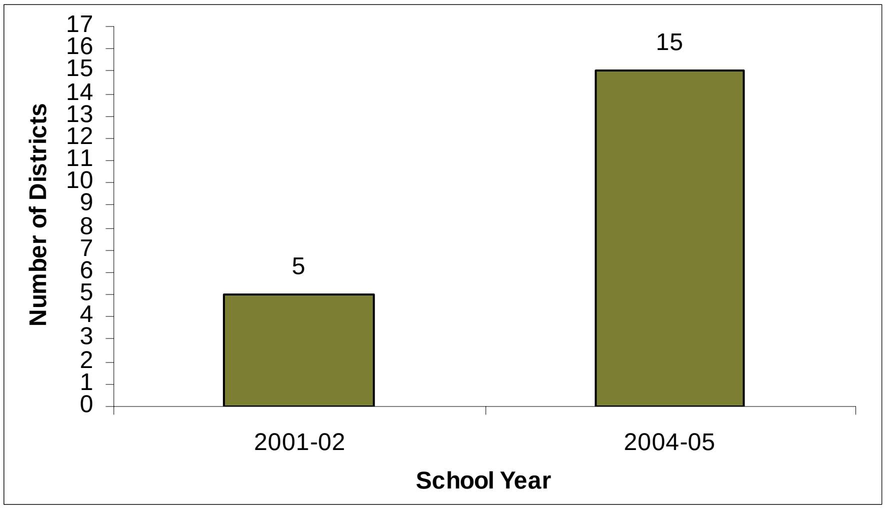
In the jurisdictions where it works, the foundation seeks to provide students and families at all income levels with the kinds of school choices that historically have been available only to affluent families. The data in Figure 11 show the number of districts that met or exceeded the foundation's 30% target for high schools of choice in 2001-02, the year in which the foundation began working in high schools, and the number of districts that met the target by 2004-05, when the initiative was 4 years into its work. Counted as schools of choice for the 17 districts and the two school years are high schools in which students could enroll from anywhere in their districts and high schools with neighborhood attendance zones that provided students with choices between two or more small, focused schools located in the large school building.[42,43] These counts do not

[42] Included in the 2004-05 counts are schools of choice that were created or redesigned with the foundation's support and schools of choice without foundation funding.
[43] Districts use various terms to describe the small, focused schools that operate in large school buildings. In many districts, they are called small learning communities; in others, they are called small autonomous schools or academies.

include schools of choice that had selective admission policies or that served special student populations, such as over-age students, court-involved students, and students with special learning needs.

The number of districts meeting the foundation's target for high schools of choice tripled between 2001-02 and 2004-05 (Figure 11).

Figure 11. Numbers of Districts Meeting the Foundation's Target for High Schools of Choice in 2001-02 and 2004-05



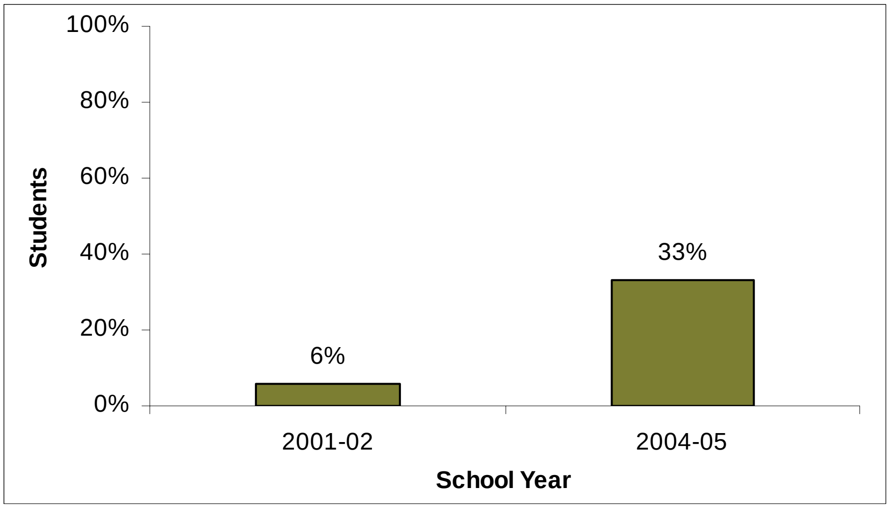
Data sources. Extant data from the 17 districts

The data show that few of the districts that we studied had the building blocks for school choice in place when the foundation began working in high schools. In 2001-02, only 5 of the 17 districts (almost 30%) provided choice to students and families in 30% or more of the districts' high schools. Across the 17 districts in 2001-02, on average, 6% of students attended high schools of choice (Figure 12).

In 2004-05, 15 of the 17 districts (almost 90%) met or exceeded the foundation's 30% target for high schools of choice (Figure 12). Included among these districts were school systems with small and large high school populations and varied histories of school choice. On average, 33% of students attended high schools of choice across the 17 districts in 2004-05. Five of these districts had more than half of their students attending high schools of choice in 2004-05.

The average percentage of students attending high schools of choice across the 17 districts increased fivefold from 6% in 2001-02 to 33% in 2004-05 (Figure 12).

Figure 12. Percentage of Students Attending High Schools of Choice in 2001-02 and 2004-05 in 17 Districts Where the Foundation Has Helped Create and Redesign High Schools



Data sources. Extant data from the 17 districts

Although these data do not establish causal connections between the foundation's work in these districts and the increase in public school options there, they do document trends that are consonant with the foundation's advocacy and contemporaneous with its investments.

Students and families in the 17 districts could choose from among several small, focused schools in large high school buildings or new free-standing high schools that represent different educational philosophies, curricular emphases, and instructional models. Across the 17 districts, almost 30% of the high school students attended schools that took theme-based or other innovative approaches to curriculum and instruction.

The foundation encourages its districts to build portfolios of schooling options so that students can choose the campuses and educational models that best meet their interests and needs. As already described, the foundation envisions a mix of traditional, theme-based, and other innovative schools in its partner districts. As previously noted, schools with traditional curricula offer students a comprehensive program with no specialized area of focus. Theme-based schools organize instruction around curricular themes, such as science and technology or social justice. Other innovative schools build their programs around particular instructional models, such as project-based learning, or they follow particular educational philosophies, such as open schooling (see Exhibit 6 for two examples).

Exhibit 6. Examples of Choice Among Educational Models

Porter High School does not use traditional instructional methods. Instead, it engages students in self-directed, project-based learning. Students negotiate with their faculty advisors in defining the topics their projects will address, the methods students will use to do their work, and the products they will create to demonstrate mastery. Project topics are as broad as the students' interests. For example, in 2004-05, one student completed an art project for which he created a portfolio of drawings and a binder of information about different drawing techniques. Other students did research on statewide environmental protection policies or wrote short novels. Students are responsible for planning their work, managing their time, and evaluating whether their learning goals and standards were met.

A contrasting example is a school that is organized around a curricular theme. The mission of City High School is to provide Puerto Rican, other Hispanic/Latino, and inner-city youth in general with educational experiences that close the "digital divide" between minority and nonminority students and expose students to careers in technology. The school has traditional high-school-level courses and a comprehensive leadership structure, yet the courses and leadership decisions build on the technology focus of the school. A mathematics class, for example, uses statistics related to computer sales and computer access as examples for practice and discussion. In addition, some of the school's more distinctive curricular offerings focus on technology. The school originally planned to prepare students for technology certification programs, but it is now working with a local college to broaden its information technology offerings.

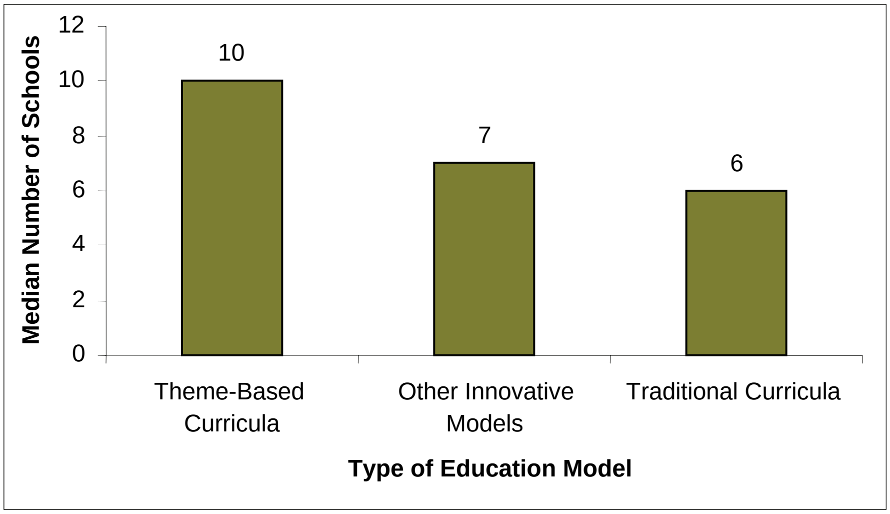
Note. All high school names in this exhibit are pseudonyms.

Figure 13 shows the median numbers of high schools of choice in the 17 districts that offered theme-based curricula, other innovative models, or traditional educational models in 2004-05.[44]

[44] It is important to point out that there were some notable differences between districts in the compositions of their portfolios, partly because their student enrollments were so variable. Four of these districts had five or fewer high schools in 2004-05, one had almost 100 high schools, and one had more than 250 high schools. Figures 13 and 14 present median values and represent reasonably well the more typical districts in this group of 17.

In 2004-05, high schools of choice in the 17 districts offered theme-based curricula, other innovative instructional models, and traditional curricula (Figure 13).

Figure 13. Median Numbers of High Schools of Choice with Theme-Based Curricula, Other Innovative Models, and Traditional Curricula in 17 Districts



Data sources. Extant data from the 17 districts

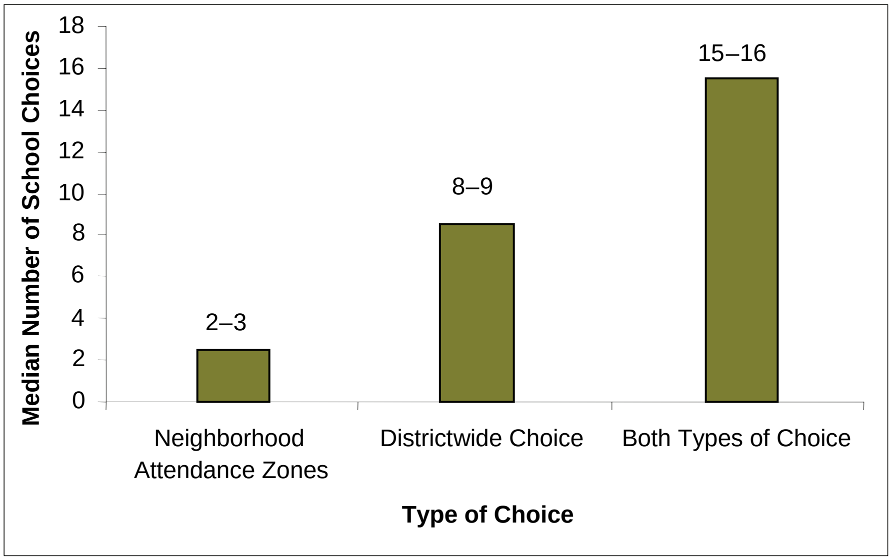
Across the 17 districts in 2004-05, a median of 10 schools or small learning communities (SLCs) (in redesigned high schools) organized instruction around curricular themes, 7 schools of choice built programs around other innovative educational models, and 6 had traditional curricula in place. Across the 17 districts in 2004-05, on average, 28% of students attended high schools of choice with theme-based or other innovative educational programs.

School choice means little if students' choices are in fact limited. Data on the numbers of choices available to students in 2004-05 are therefore important, since students are more likely to find programs that speak to their needs, interests, and future plans in districts with many available choices. In efforts to maximize student choices, almost all of the districts in the evaluation created school portfolios that included one or more high schools with districtwide enrollments. Available options included free-standing schools of choice and large high schools that were redesigned to house two or more small focused schools or SLCs. Several of these districts also provided students with choices among the SLCs in their neighborhood high schools. However, not all districts gave students the option to select a school from anywhere within the district. In one district, students had choices within redesigned high schools in their neighborhood attendance zones but not outside them.

Figure 14 provides data on the numbers of choices that were available to students in a typical district in 2004-05.

In 2004-05, students in districts that provided districtwide choice could select from among 8 or 9 different schools of choice. Students in districts that supplemented districtwide choice with choices among small focused schools located in neighborhood high schools had 15 or 16 available schooling options (Figure 14).

Figure 14. Median Numbers of School Choices Available to Students in 17 Districts That Provided Various Types of Choice in 2004-05



Data sources. Extant data from the 17 districts

In 2004-05, students had a median of 2 to 3 choices in districts where they could choose from among SLCs within their neighborhood high schools but not outside their neighborhood attendance zones. They had 8 to 9 choices in districts where students could choose high schools or SLCs located anywhere in their districts and 15 to 16 choices in districts offering both neighborhood and districtwide choice. In this last group of districts, students could choose from among free-standing schools of choice, SLCs in their neighborhood high schools, and SLCs outside their neighborhood attendance zones.

As an illustration of one district's approach, Exhibit 7 provides examples of the types of choices that were available to students in 2004-05 in the Oakland Unified School District.

Exhibit 7. Examples of Students' Choices in Oakland Unified School District[45]

Oakland has a choice policy, called Options, in which entering freshmen can choose from school programs throughout the district. These choices include small learning communities within three redesigned high schools, several new small high schools, and conventional high schools. Oakland's vision extends beyond school choice to encompass school diversity. The district's goal is to offer a wide range of options—including theme-based schools and schools with different pedagogical models—where students can work at internships, engage in Expeditionary Learning, or become a cadet at a military academy. The following are examples of choices students had available in 2004-05.

Choices Among Small Learning Communities in Redesigned High Schools

Castlemont Community of Small Schools:

- **Leadership Preparatory High School:** Leadership Preparatory focuses on literature, writing, culture, philosophy, and media arts. Students can take Advanced Placement classes or can enroll in courses at local colleges and universities.
- **Business and Information Technology School:** This school prepares students for college and the high-tech workforce. It uses project-based learning strategies and incorporates computers and other technology into courses.
- **East Oakland School of the Arts:** This school combines a focus on the arts with a strong academic curriculum, preparing students for postsecondary education or a professional career.

Fremont Federation of High Schools:

- **College Preparatory and Architecture Academy:** This academy offers a college preparatory curriculum and electives focusing on architecture.
- **Mandela High School:** Mandela High is college preparatory with a focus on critical literacy and social justice. The school's theme is "Bridging Multiple Worlds."
- **Media Academy:** This academy focuses students on careers in newspaper, magazine, radio, and television journalism. The school offers a bilingual language arts curriculum.
- **Paul Robeson College Preparatory School of Visual and Performing Arts:** This school exposes students to cultural diversity through the arts while delivering a strong academic curriculum. Students have opportunities for involvement with the school and surrounding community.

McClymonds High School:

- **Business Entrepreneurship School of Technology:** This school encourages the development of lifelong learners, using project-based and real-life learning activities. Students are prepared to either enter college or embark directly on a career.
- **The Experience, Excellence, Community, Empowerment, and Leadership School:** This school focuses on service and activism and houses two academies: the Stories, Media, and Arts Academy and the Law and Government Academy.

Choices Among New Schools

- **East Bay Conservation Corps (EBCC) Charter School:** The Corpsmember Program High School Level of the EBCC Charter School allows EBCC Corpsmembers, ages 17 to 24, to complete a high school diploma or GED. The Corpsmember Program includes both on-the-job training in environmental improvement projects and formal classroom instruction.

[45] From the Oakland Unified School District Web site (http://webportal.ousd.k12.ca.us/index.aspx).

Exhibit 7. Examples of Students' Choices in Oakland Unified School District (*continued*)

- **East Oakland Community High School:** This school is based on principles of youth development. Using a project-based learning approach, students are empowered as producers, not just consumers, of culture.
- **Far West:** This school focuses on hands-on learning and critical-thinking skills. Staff work to build relationships with parents and the community to assist in educating students. The school has a partnership with a nearby arts college.
- **LPS Oakland:** This school is a charter school that operates out of the Castlemont facility. The school has advisories, and the focus is on leadership skills.
- **Life Academy:** Life Academy has a health and bioscience theme. Students are prepared for college while learning about opportunities in medicine, mental health, biotechnology, and science.
- **Lighthouse Community Charter School:** This K-12 school follows nine Guiding Principles, which include meeting the needs of the whole child and articulating high expectations. An individualized learning plan is created for every student. The school is affiliated with Expeditionary Learning Outward Bound.
- **Lionel Wilson College Preparatory Academy:** This school is part of the Aspire Public Schools network, which focuses on high standards and personalized learning. The school is college preparatory, with all members of its first graduating class accepted into 4-year colleges.
- **Merritt Middle College High School:** This school, located on a community college campus, works with the college to provide education for students for whom large, traditional high schools did not work.
- **MetWest:** MetWest is a Big Picture school, focused on "one student at a time" learning. Teachers serve as advisors to students, who spend 2 days a week at internships and 3 days at school in academic study and working on projects.
- **Oakland Military Institute:** Students at this school are called "cadets." They are expected to excel academically and personally. The school uses a military framework to achieve its goals.
- **Oakland Unity High School:** Unity High is a college-preparatory school that stresses support for students through partnerships between the school, families, and the community. Students are offered leadership opportunities, and each one is assigned a mentor for informal counseling.
- **Oasis High School:** Oasis focuses on students who want to complete high school but have had negative experiences in the past with academics. The school offers small classes and one-on-one advising for students.
- **University Preparatory Charter Academy:** The U Prep curriculum is linked to courses at local community colleges. U Prep works closely with these colleges, as well as with parents, to support every student. U Prep has a longer school day than normal, as well as an 11-month school year.

District Partnerships to Manage School Reform

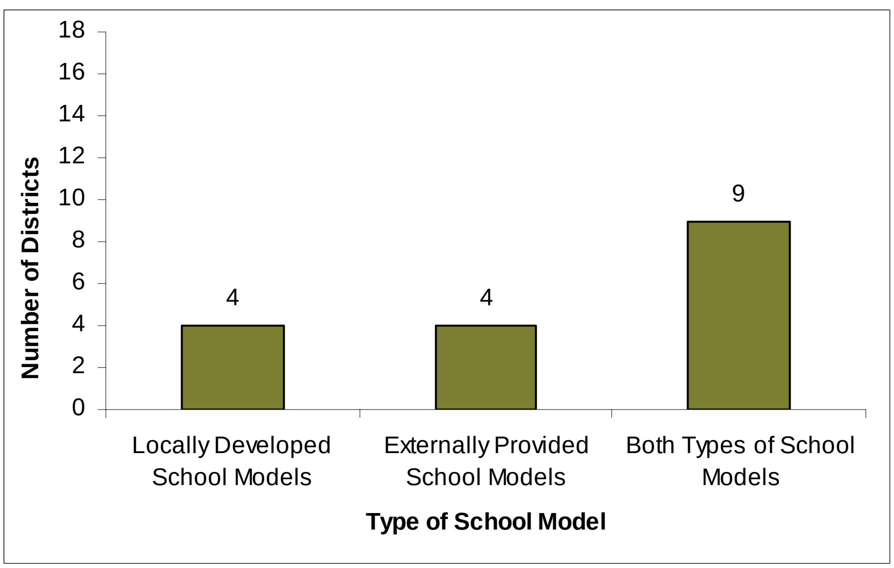
In some of the school systems that we studied, district and school staff developed their own school models, designing curriculum and instruction for their schools and setting up the academic structures for them. In others, district and school staff adopted school models that were supported by school franchises and other external providers. Many districts put in place a mix of locally developed and externally provided school models.

The foundation works with districts in a variety of ways. In most cases, it provides funds to organizations that use a range of strategies to support reform. These organizations provide funding to districts and schools, help with charters and facilities, provide back-office support and technical assistance to districts and schools, and, in some cases, help teachers and other school leaders refine their thinking about the curricula, instructional methods, and academic structures that define new free-standing high schools or SLCs. In other cases, these reform organizations help faculty implement teaching and learning models that are in use elsewhere in the country and have proven to be effective. Such reform organizations as the Center for Collaborative Education and the Center for School Change help district and school personnel develop school models that meet local needs and interests. They also provide support as schools implement locally developed models. In contrast, such organizations as The Big Picture Company and High Tech High help schools implement school models for which they have demonstration sites, provide resource materials, offer teacher professional development, and provide networking opportunities. With foundation support, some districts allow for a mix of locally developed and externally provided school models in their high schools.

Figure 15 provides data on the types of school models that were in use in the 17 districts in 2004-05. It shows the number of districts in which schools created their own school models, the number of districts in which new schools adopted externally developed school models, and the number of districts that supported schools as faculty developed their own or adopted externally provided school models.

In 2004-05, more than half of the districts in our evaluation created new and redesigned high schools using both locally developed and externally provided school models (Figure 15).

Figure 15. Numbers of Districts That Created New and Redesigned High Schools Using Various Types of School Models in 2004-05



Data source. Extant and/or interview data from the 17 districts

In 2004-05, schools in 9 of the 17 districts that we studied adopted a mix of locally developed and externally provided models in their new and redesigned high schools.

Other districts implemented either locally developed or externally provided school models. Schools in 4 of the 17 districts developed and implemented locally derived models. One of the four districts developed SLCs from scratch by inviting teachers and other school leaders from district high schools to submit proposals for curriculum and instruction in the SLCs they hoped to create. District officials encouraged faculty to think expansively about curricular themes and instructional models. The resulting small schools reflected the passions of faculty and the needs and interests of the students.

New schools in four districts adopted solely externally provided school models. Seeking an alternative to its existing conventional high schools, one of these school systems invited The Big Picture Company to develop a school for the district. Responding to this invitation, Big Picture provided the academic model and curriculum for the school. Although the school is operated and managed by the district, it maintains close contact with the Big Picture network through coaching, weekly networking calls, and faculty participation in Big Picture conferences.

One of the nine districts that adopted a mix of locally developed and externally provided models used an "incubation" process to help faculty develop their concepts for small learning communities over a year of planning. In addition, the district invited national reform organizations, including the Knowledge Is Power Program (KIPP), Expeditionary Learning Outward Bound, and The Big Picture Company, to develop schools of choice for them. Exhibit 8 describes the work of reform organizations that help build local school models and provide external models.

Exhibit 8. Examples of Reform Organizations That Help Build Local School Models and Provide External Models

The Center for Collaborative Education (CCE) is a reform organization that provides districts and schools with five types of services: visits to an established small school, technical assistance and coaching, summer institutes and school year meetings, funding, and access to a pool of new small-school leaders (http://www.nessn.org/nessnconcept.html#offers). CCE staff work closely with the districts they serve, generally spending 2 to 3 days per week in the district. CCE does not require schools to follow a particular pedagogical model or use a specific curriculum. However, it espouses a set of principles which it recommends that schools also embrace (http://www.nessn.org/nessnconcept.html#criteria). These principles describe (a) schools where students who work to use their minds well and become responsible members of a democratic society, students master a limited number of essential skills and areas of knowledge, take responsibility for learning, and use exhibitions and portfolios to demonstrate mastery; (b) schools in which high expectations and equity and access are key, student-teacher ratios are low, and teachers and students know each other well, governance is democratic, and students take responsibility for learning, they use exhibitions and portfolios to demonstrate mastery, trust and respect are pervasive; and (c), teachers who create professional collaborative communities, and governance is democratic. CCE also recommends that schools operate as schools of choice with enrollments between 50 and 400 students and commit to securing long-term facilities and allowing small-school autonomy.

Exhibit 8. Examples of Reform Organizations That Help Build Local School Models and Provide External Models *(continued)*

The Big Picture Company creates high schools of choice based on the model of the Metropolitan Regional Technical and Career Center (The Met) in Providence, Rhode Island. Big Picture schools educate "one student at a time" through individualized curricula based on student interests and passions, authentic learning through internships with mentors, and public exhibitions of student work (http://www.bigpicture.org/aboutus/philosophy.htm). Big Picture Schools use the same language and engage in the same practices; they enroll fewer than 150 students,; and teachers act as advisors. Big Picture begins its work with districts by selecting school principals and supplementing the district teacher hiring process, which it often finds is inadequate for its needs. Big Picture works with districts during TYBO (The Year Before Opening) principal training, assessing what the principals will need in their particular district and helping to ensure that they get it. Further, it works with TYBO principals on design implementation plans and figuring out how state and district requirements and college admission requirements fit with the Big Picture model. When schools are up and running, professional development takes place at staff meetings, regular staff retreats, and summer conferences. Staff and advisors participate in Big Picture professional development activities, including the annual "Big Bang" conference and videoconference workshops.

Challenges and Supports to Reform Efforts

In all of the districts we studied, there was evidence of some of the challenges and supports the foundation recommends, including college-ready missions; effective governance; strong accountability; community engagement; and supports for curricula, teachers, schools, and students and families. None of the districts had all the recommended challenges and supports in place.

Foundation officials stress the difficulty and complexity of school improvement in underperforming school districts. They explain that local reformers must navigate "layers of local, state, and federal regulations; restrictive employment agreements; antiquated management systems; and, perhaps most damaging, a culture of differential expectations, compliance, or helplessness" (Bill & Melinda Gates Foundation, 2005a, p. 4). To support reformers' work, the foundation has asked districts to commit to the development of challenging and supportive operating environments. The foundation contends that supportive operating environments challenge schools to perform at high levels by doing the following:

- Adopting common standards that prepare all students for postsecondary education, work, and citizenship
- Focusing on results and equity and empowering improvement with transparency, measurement, and responsiveness, as well as alignment with state goals and policies
- Implementing transparent performance management systems with steps of progressive intervention that provide support for students, staff, school, and systems

- Engaging parents, citizens, and business and civic leaders in making quality education a priority

The foundation also asks districts to put in place systems to support school improvement, including:

- Learning expectations that provide a spine for instructional materials, diagnostic assessments, ramp-up supports for students, and teacher development activities
- Strong learning and support networks for schools with dollars that follow students and reflect student needs, budgets that allow school-based decision-making; needed core services, and optional purchased services
- The identification and development of instructional leaders, district recruiting, school/network-based hiring, 3-year induction with ongoing, job-embedded development in a professional learning community, and compensation that reflects ability, performance, and responsibility
- A student support system with an advocate for every student to ensure appropriate guidance and academic support, connection to family services, and frequent and specific performance feedback against clear expectations

Although none of the 17 districts on which this chapter focuses have put all of the recommended challenges and supports in place, all of the districts have some of these characteristics. What follows is a discussion of challenges and supports in these districts, including examples of barriers to the implementation of supportive operating environments.[46]

Challenges and Supports. Faculty members' efforts to increase rigor in their classrooms were discussed in chapter 3 of this report, but note should also be taken of work done at the district level to promote rigor. The 17 districts have taken a number of steps to challenge their high schools to reach for high standards and have provided supports for the schools to do so.

Like many across the county, most of the districts we studied had content and performance standards in place in 2004-05. Some also offered diagnostic assessments and benchmark tests. Several jurisdictions provided coaching and other ramp-up supports for students who struggled on diagnostic tests. One district offered quarterly benchmark tests and provided teachers with the data they needed to make classroom-level instructional adjustments and to tailor instruction to individual student's needs.

Some districts promoted high standards by brokering relationships between high school faculty and local colleges and universities. According to district staff, these partnerships addressed several goals. Among the goals were improving high school students' preparation for college application and admission, expanding the pool of tutors for

[46] Though the principal survey was not designed explicitly to address the challenges and supports the foundation recommends, we used principal survey data from schools in the 17 districts to measure some of these constructs. We could find no systematic relationships between the measure we created and the other variables reported in this chapter.

struggling high school students, providing high school faculty with the help of preservice teachers, and expanding opportunities for university-based teacher professional development.

One district supported faculty in its redesigned high schools by providing external coaches for teacher teams from the redesigned schools. Coaches worked with SLC teams on methods for teaching mixed-ability student groups, improving teaching practices, and integrating new teaching tools. In the second year, the district increased the number of coaches and hired a full-time staff member to coordinate coaching and the teacher teams.

As already mentioned, some districts challenged their redesigned schools by encouraging faculty to adopt proven school models and to use curricular approaches that are widely known to be successful. District officials asked faculty to seek the help of the reformers who support these school designs and take advantage of their instructional resources and professional development.

Concerned about students' weak performance in mathematics when they enter high school, one district offered professional development to its middle and high school teachers to create seamless transitions in mathematics instruction between the two levels of schooling. The district sponsored alignment workshops through which faculty in different grades discussed mathematics curricula and instruction in the years prior to and at the beginning of high school. They also talked about the strengths and needs of students making the transition.

Several of the districts took alignment strategies one step further by exploring opportunities to add middle school grades to new high schools. One district already had a new high school that included middle school students; two had plans to add middle schools; and another district began work with a local university to create a K-12 school. Officials in these districts described the importance of combating low levels of preparedness at high school entry.

One of the 17 districts is raising expectations by eliminating middle and high school tracking. Others are following suit by expanding access to honors courses, Advanced Placement classes, and the AVID (Advancement Via Individual Determination) program.

In 2004-05, districts used a number of strategies to engage their communities in high school improvement. Several districts held community forums to generate parental involvement in school reform. Districts involved stakeholders in planning improvements and evaluating reforms, supporting school partnerships with business and government agencies, hiring family-involvement coordinators or public-engagement coordinators, and including public engagement in school performance management plans. In many districts, district staff and faculty consider outside supports and relationships integral to their work.

One district partnered with the Annenberg Institute for School Reform at Brown University to create a review process to help the district evaluate its capacity to support high academic performance for all students in all schools. Participating in the review are community members, parents, teachers, administrators, students, board members, and Annenberg staff members. In 2004-05, this group worked together to identify high-priority issues in the district, examine quantitative and qualitative data related to these issues, and develop action plans for school improvement.

Barriers to Reform. Although there was evidence in 2004-05 that some of the 17 districts had implemented several of the recommended challenges and supports that the foundation encourages, some districts had also implemented policies and practices that created barriers to reformers' work. Financial barriers were perhaps the most common result of these district practices. Only a small number of the districts we studied, for instance, provided start-up funding to their new and redesigned schools. New schools frequently had low enrollments and insufficient operating funds under their state funding formulas. Many of the schools struggled financially and were targets for community concern about possible "encroachment" on districts' general funds. In one such district, a newly appointed superintendent expressed concern about this situation and vowed to support current schools but not open others that could not sustain themselves financially.

More general budget cuts in several districts also posed problems. In one district, budget cuts resulted in staff reductions in the district's redesigned high schools. These reductions required the remaining diminished teaching staff to "cross over" and teach classes in more than one small school. School staff members considered this development a breakdown in the borders between the small schools and a step toward the destruction of the individual SLC identities. Staff in several of these districts predicted future financial difficulty for their SLCs and the need to find sources of supplementary support.

In one district, officials mandated cuts to teachers' common planning time. Teachers in another district reportedly worried about district commitment to school improvement and to their school. In this and other districts, district hiring procedures also limited schools' ability to attract teachers who bought into their educational models.

Sustaining School Reform

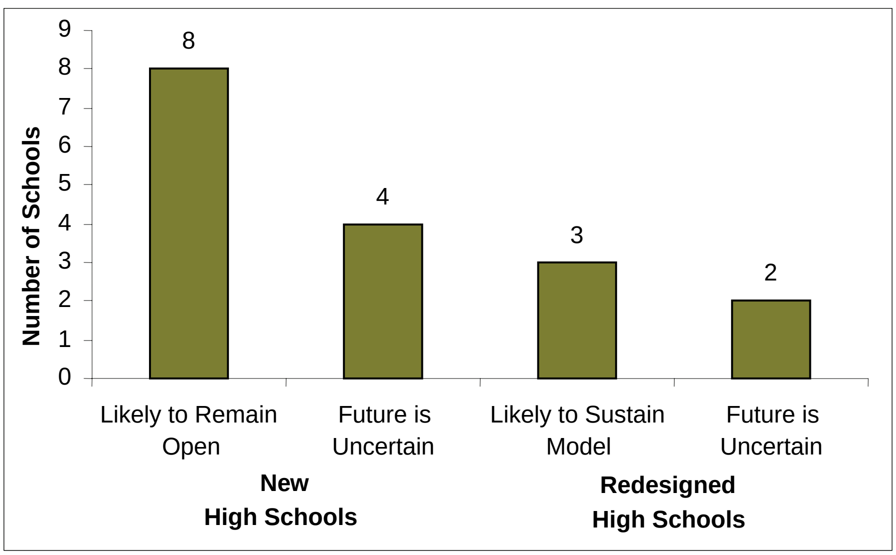
By 2004-05, faculty at 8 of the 12 third- and fourth-year new schools in our evaluation felt confident that their schools would survive in the long run. Additionally, faculty at three of the five third- and fourth-year redesigned schools we visited predicted that the instructional models that defined their small learning communities would remain intact in 2005-06. In the other new and redesigned schools, some faculty members were uncertain about the futures of their programs.

The finding that each of the 17 districts has in place some but not all of the recommended supports has important implications. Put simply, in the absence of some of the needed supports, the futures of the foundation-supported 12 new schools and 5 redesigned schools that we studied are uncertain. This section describes what the future may hold for some of these schools.

When interviewed in 2004-05, faculty at 8 of the 12 third- and fourth-year new schools in our evaluation felt confident that their schools would remain open in 2005-06 (Figure 16). These faculty predicted that their schools were here to stay and that their instructional models were likely to endure. In contrast, faculty in the other four new schools said that they were uncertain about the future of their schools. These faculty members indicated that some of the barriers discussed above were at the root of the doubts about sustainability. Faculty in one school, for example, reported waging an uphill battle with the district to improve student test scores. The school had failed to make adequate yearly progress (AYP) on the state assessments under NCLB. The resulting need to focus on test preparation purportedly worked at cross-purposes with the school's educational philosophy and instructional design.

In 2004-05, school faculty in two thirds of the new high schools and just over half of the redesigned schools predicted their programs would remain intact in 2005-06 (Figure 16).

Figure 16. Predicted Sustainability of the Foundation-Supported New High Schools and Small Learning Communities in Redesigned High Schools



Data sources. Interviews with faculty at the 12 new and 5 redesigned high schools

Officials at another of the new schools with an admittedly uncertain future were very blunt about needing ongoing support once foundation funding ends. The school leader explained that his faculty was putting "energy into areas that we wouldn't typically have to: grantwriting, being creative with our programming, moving money around, taking longer for goals to actually happen. I mean, money makes this thing tick."

Because redesigned high schools enroll most of the students who previously populated their parent large high schools, they are less likely to close their doors. They are less of a novel experiment and more of a reworking of an existing institution. Nonetheless, redesigned schools face the challenge of sustaining their instructional models and SLC structures. In our 2004-05 interviews with faculty at five redesigned high schools in their fourth year of reform, faculty at three said that they were confident that their schools would sustain the reforms begun with the redesign. Staff at the two other redesigned schools, however, were less confident that their schools' programs would continue.

Whereas teacher turnover in urban high schools across the country averages 18% per year, 30% or more of the teaching staff in the foundation-supported new schools left after each of the schools' first 3 years of operation.

The high turnover rates among faculty at the new schools pose additional threats to school sustainability. Analyses show that the new schools in the studied districts experienced substantial teacher turnover in their initial years.[47] Whereas nationally reported teacher turnover rates in high-poverty schools average about 18% per year (Provasnik & Dorfman, 2005), it was typical for 30% of the faculty in the studied foundation-supported new schools to leave after the schools' inaugural year. Similarly, approximately 30% of second-year faculty left at the end of the second operational school year. Between the third and fourth school years, on average, 40% of the faculty left their schools. During these same years, new schools were adding grades and students. Consequently, in addition to replacing departing staff, the schools had to increase their faculty numbers by 10% to 15% to accommodate growing enrollments.[48]

The faculty members with whom we spoke reported several negative effects of turnover in their ranks. These complaints included the difficulty that turnover created in implementing a strong vision, the recurring need to train new teachers, the resulting loss of opportunity for advanced teacher professional development, and a general decay in teacher professional community. Teacher turnover was particularly high among the ranks of mathematics and science faculty.

Teacher turnover was not always a negative phenomenon, however. In fact, in some instances, turnover presented a growth opportunity for schools. Interviewees at some schools suggested that when staff left, their schools were able to hire new staff who represented a better fit with the school model.

According to faculty, the AYP and "highly qualified teacher" requirements of NCLB compound hiring problems in these schools. Some of the new schools were staffed by as few as four teachers, and finding certified faculty who were knowledgeable about innovative learning models and who were willing to work collaboratively was

[47] We did not have teacher rosters for consecutive years for the redesigned high schools. We had faculty lists from the year schools were planning for redesign and then 2 years later in the schools' SLCs. It was common in districts with redesigning high schools for some faculty to change buildings between the planning and first redesign years. Comparisons between teacher rosters for the planning and second redesign years would have overestimated teacher turnover rates in redesigned high schools.

[48] Teacher rosters were collected from 12 new schools that were in the third or fourth year of operation in 2004-05. School rosters were compared across academic years, and the numbers of faculty who were retained and newly hired and those who left in 2001-02, 2002-03, 2003-04, and 2004-05 were cataloged. Data were combined so that the percentages of teachers who stayed on staff, came on board, and left as schools moved from their first to second, second to third, and third to fourth years could be calculated.

challenging. This challenge was compounded at schools with project-based instructional models, in which teachers work as advisors to students in all subject areas, and at schools with thematic curricula, where instruction is generally multidisciplinary. Finding teachers with the proper certifications for these school models was particularly problematic.

Implications for the Initiative

Foundation officials are painting a new landscape for secondary education, helping to develop school systems that give all students the tools for success. In most of the studied districts, the foundation has achieved its goal of offering choice in at least 30% of the communities' high schools. Students and families have more educational choices than they did before the foundation began its work, and they have choices among numerous free-standing schools and small learning communities. New options include schools with traditional, theme-based, and other innovative instructional models. In most of the studied districts, the average student had eight or more different schools or small learning communities from which to choose. In districts where students' primary choices were among SLCs within their neighborhood schools, the number of choices was smaller. It is important to note that it is too early in the life of the initiative to make judgments about the quality of these school options.

Many of these schools of choice are likely to be sustained beyond the grant term. Staff members suggest that many of the schools that the foundation opened are likely to remain open and maintain their instructional designs. The instructional models of small learning communities in many of the redesigned high schools are also likely to remain intact.

School districts challenged their schools in one or more of the following areas: organizational mission, governance, accountability frameworks, and community engagement. They provided some of the foundation's recommended supports for curricula, schools, teachers, and students. None of the districts had all of the recommended challenges and supports in place. Despite the many good components of district portfolios, the foundation, its grantees, and the partner school districts should be concerned about these missing features and the resulting lessons learned from the schools that are not as optimistic about their futures. There are clearly areas that invite more attention from all participants in the initiative.

There is evidence that some districts hosting foundation-supported schools have not bought into the foundation's particular reform vision. Since district buy-in is an ingredient for success, it may behoove the initiative to engage in more vigorous outreach to districts, and the foundation may wish to exercise greater financial leverage in the districts than it historically has. To strengthen the alignment between its vision for district support and district practice, the foundation may want to revisit its memoranda of understanding with districts to clarify requirements and to make stronger stipulations about district commitment to the recommended challenges and supports.

Commentary from many school leaders underscores the degree to which schools are vulnerable to the high turnover rate among teachers. This vulnerability is acute in many new schools, which are generally small, and it is most acute in mathematics, the discipline in which districts have had difficulty finding adequate numbers of certified teachers. As the initiative moves forward, the foundation and its reform partners may wish to build teacher recruitment and retention strategies into their reform efforts.

Teacher recruitment and retention are serious problems in many districts that serve disadvantaged students; the issues are not new to initiative participants. Innovation will lie in the combinations of approaches that participants create to address the problem. For instance, reforming schools may want to consider tapping into state-approved alternate certification procedures for teachers in disciplines in which there are shortages of certified faculty. They also may wish to target professionals from key fields who are either retiring or considering career changes. Mentorship programs and other retention strategies may be particularly important for teachers in reforming schools. In the future, the foundation may even want to make its reform efforts contingent on a district's providing support for teachers—where support includes extra financial and peer support systems, as well as professional development opportunities.

Finally, the lessons that individual schools have learned about strategies for sustainability should become part of the initiative's lore. Publications and professional development efforts targeting school officials should make use of these lessons so that current and future school leaders can prepare themselves for the many obstacles that confront school reform efforts.

Chapter 5. Themes and Recommendations

We conclude this report by highlighting some of the major themes emerging from both the specific findings described in chapters 2 through 4 and findings from earlier reports in this series (AIR/SRI, 2003; AIR/SRI, 2004; Mitchell et al., 2005; Rhodes et al., 2005; Shear et al., 2005). These themes provide the backdrop for a set of recommendations proposed to the Bill & Melinda Gates Foundation, as well as to other organizations promoting large-scale high school reform.

Themes Emerging From the Evaluation

We begin by considering what has been learned with respect to both the process and the outcomes associated with starting new high schools as a way to serve students better. We then consider the high school redesign strategy, how the foundation structures its work with organizations receiving grants, and the prospects for system change.

- **The new high schools created with foundation support are fostering better outcomes for students in terms of engagement, attendance, grade-to-grade progression, and the quality of student work in English/language arts.** Although the new schools have struggled with challenges in terms of staffing, financing, curriculum development, and setting behavioral expectations (AIR/SRI, 2003), collectively they clearly provide a qualitatively different—and in many ways more positive—high school experience. At the same time, these schools are enrolling students who enter high school behind their peers in terms of prior measured achievement, and the schools' record with respect to academic progress, as captured by standardized test scores, has been uneven (Rhodes et al., 2005).

  When considering these outcomes, one should be aware that the foundation's criteria for awarding high school grants have evolved since the initiative started in 2000. As noted in chapter 1, the foundation originally awarded grants to individuals and organizations promoting the attributes of effective schools. More recently, however, the foundation has placed less emphasis on educational philosophy and greater emphasis on grantees' having a well-specified intervention model and a track record of raising student achievement. The current evaluation study did not include enough schools resulting from the most recent grants to assess the impact of this change in criteria. Thus, at this stage we cannot say whether the later grants are producing schools with student outcomes different from those experienced at the first schools created with foundation support.

  Our evaluation of outcomes is limited also by the schools' relatively short histories. There were not enough new schools with graduating classes to assess the initiative's success in achieving the foundation's central, longer term goals—

that is, improving rates of underserved students' high school graduation, college admission, and college completion. The attention that many of these new schools devote to individual students' needs and interests and the willingness of school staff to take extraordinary steps to keep their students in school provide some grounds for optimism. Nonetheless, more longitudinal data are needed to test the success of the initiative with respect to producing college-ready graduates. Unfortunately, access to the types of data needed for this kind of analysis is hindered by the fact that many of the foundation-supported schools and their districts lack a system for keeping track of individual students and their academic outcomes once a student has left a school (as a transfer, a dropout, or a graduate).

- **The new high schools have not had the same kind of success in terms of standardized achievement test scores or in terms of students' class work in mathematics.** Mathematics learning appears weak, whether measured through standardized tests or through expert teachers' ratings of student work. Students entering the new schools, on average, are behind their peers at other district schools in terms of prior achievement. Because mathematics as a content domain is more cumulative and sequential than many other subjects, staff at new schools find it particularly difficult to deal with heterogeneous mathematics classes in which some students are ready for trigonometry while others are still struggling with manipulating fractions. The new small high schools have had difficulty recruiting qualified teachers of mathematics, and many of the teachers that they do hire lack substantial mathematics teaching experience. Some of the schools also report a lack of well-established instructional approaches for high school mathematics that are compatible with their student-centered instructional philosophies.
- **New high schools face a number of challenges to their continued program development and operation, including serious issues related to staff retention and to staff recruitment in some critical subject areas.** New high schools' recruitment difficulties are not limited to finding qualified mathematics teachers. In general, small schools face challenges in recruiting staff with the qualifications for teaching all the courses required in a college-preparatory curriculum (particularly now, when No Child Left Behind requirements for highly qualified teachers are being implemented). The challenge is heightened further because the schools seek teachers who are certified in each subject area and who also have the disposition to act as advisors and confidants for students who face significant personal and social barriers to academic success.

  Stressful work environments create an additional challenge to new schools' efforts to retain staff members. Staff at the new schools describe the long hours required to be a teacher, curriculum developer, part-time administrator, and student advisor and the strain resulting from their emotional involvement with their students' problems (AIR/SRI, 2003). In some schools, staffing appears to be stabilizing, with teachers feeling part of a cohesive team. But on the whole, teacher turnover at the new schools has been high, as many teachers conclude that the job takes too much out of them to be feasible over the long term.

At the same time, our analyses of teacher interview data suggest another kind of staffing issue: finding teachers who have the requisite faith in their students' capabilities. Even at these new schools of choice, some teachers express doubts about their students' ability to do college-preparatory work.

- **Within a 3-year time frame, positive outcomes for students appear more likely to be obtained through starting new high schools than through redesigning existing schools with poor track records.** It takes longer to redesign and change the culture of a high school than to start a new one, and the early record of student outcomes suggests that the effort to redesign schools may be of limited efficacy. In its first 2 years, redesign typically produces some improvement in terms of school climate. When results of the smaller units created from redesign are averaged, however, we have found little or no impact on attendance or student achievement. Our qualitative data suggest that attention and resources in many of the foundation-supported school redesign efforts were focused on structural changes and that issues of curriculum and instruction at these schools were postponed (Shear et al., 2005). Now, as noted in the introduction to this report, the foundation is pressing schools undergoing redesign to address core instructional issues in their first year. Outcomes in student achievement may improve correspondingly.

  Other analyses comparing outcome data for the small learning communities (SLCs) resulting from school redesign raise further issues. We have seen the potential for students and teachers to self-select into SLCs in ways that produce different levels of academic challenge and different climates in the various subunits (Evan & Song, 2006; Shear et al., 2005). Ironically, it appears that SLCs run the risk of inadvertently becoming unofficial tracking systems, with certain historically underserved groups ending up in subunits where there are fewer opportunities to engage with challenging content.

- **The diverse nonprofit organizations that received foundation grants have been supporting high schools in multiple forms; however, individual grantee organizations typically have strong capacity to support schools in some ways but limited capacity in other areas, with secondary curriculum being a common weakness.** We have seen that in some cases these organizations garner community support for new or redesigned high schools and, importantly, that they can continue to promote the school reform even if the superintendent or school board turns over. In other cases, grantee organizations have provided their schools with expert consulting on the school design process and on alternative school structures.

In a few cases, the grantee organization has brought resources and expertise in curriculum and instruction. However, the typical grantee organization has not been able to play all these roles effectively for all the schools it is supporting (AIR/SRI, 2004). As a result and to get the support they need, schools often need multiple partnerships and supporters who are well positioned politically.

- **An increase in public school choice appears to be occurring in the jurisdictions where the foundation has focused the most effort.** Because this evaluation study collected data primarily at the school level rather than the system level, our inferences about systems must be guarded. Nevertheless, we did collect data on the district contexts within which the foundation-supported schools are operating. We found that a large majority of the 17 districts in our analysis had reached the target "tipping point," with 30% or more of their high schools being schools of choice. While some districts with foundation-supported schools are satisfied making available just one or two distinctive small high schools of choice, and some are averse to profound change in the organization of most of their high schools, there are several districts where we have seen deep restructuring with a reasonable chance for lasting change. In Oakland and New York City, for example, it is not only the portfolio of available high schools that has changed but also the operation of the district and the way in which the district is structured.

Recommendations for the Initiative

On the basis of our experience examining the progress of foundation-supported new and redesigned schools over the past 4 years, we have derived a set of recommendations concerning areas of need and potential refinements to the high school reform initiative. We note that these recommendations are inspired by our findings but go beyond our research results per se. They are the opinions of the authors and not necessarily the policy of the Bill & Melinda Gates Foundation.

- **Maintain the focus on student academic outcomes and on curriculum and instruction as the means to attaining those outcomes.** Increasingly, the foundation is stressing the importance of focusing on student outcomes, with concrete targets and a system for measuring progress toward them. The benefits of this strategy are that it keeps the focus (appropriately) on students and it offers reformers latitude in approaches as long as schools and systems of schools are able to demonstrate strong outcomes. Such a strategy can be effective, however, only if there is a system in place for monitoring and tracking the student outcomes that matter. Agreement about outcome measures (e.g., proficiencies, graduation rates, college admission, demonstrated proficiency) and a system of tracking and reporting progress for every student should be prerequisites for receipt of foundation support. The foundation should not expect every school to develop and implement such a system independently, however. A centralized effort to provide tools and training is needed.

- **Exert leadership in aggregating and disseminating the most effective instructional resources for serving high-need secondary students.** High-end secondary curricula in the form of Advanced Placement and International Baccalaureate programs are available today and are being promoted in a broader range of school settings. Many students, however, enter high school with skills that are two or more grade levels below their nominal grade placement. These students need additional opportunities to learn in order to equip themselves for success in grade-level course work. As noted above, the new high schools are finding it particularly challenging to fill this need in mathematics.

  Approaches that move students to grade-level proficiency (and beyond) and that can be used in classes where students have a wide range of skill levels are needed. The decentralized approach to curriculum and instruction that the initiative (and, some would say, the U.S. education system as a whole) has used to date has not been effective in helping students at foundation-supported high schools achieve at high levels in mathematics. The foundation has the opportunity to exercise leadership by setting up a collection of those secondary education instructional materials and approaches that have the best evidence of effectiveness with traditionally underserved groups. It also has the opportunity to test the effectiveness of unproven but promising approaches and to disseminate the lessons learned from the more efficacious ones. Setting up the collection would entail convening a panel of experts to establish criteria for high-quality materials and applying the criteria to identify the best available instructional resources. The panel and the foundation could invite submissions of materials along with evidence of effectiveness. Those approaches and materials approved by the panel could be collected and synthesized in ways that help schools and teacher training programs identify high-quality instructional resources that are (a) appropriate for their target student groups and (b) aligned with the schools' and programs' instructional philosophies. Those schools and districts receiving foundation support could be urged to draw on this instructional resource library and be given technical assistance if they choose to implement one of the recommended instructional approaches. Once such approaches are implemented, the foundation could commission rigorous studies to examine their effectiveness.

- **Explore alternative models for supporting teachers in schools with high-need students.** Given a strong system of personal and academic supports, many more students can succeed in completing high school ready for college or productive work. Providing that range of supports, however, is labor intensive, and we have seen the toll it can take on dedicated teachers in small high schools. We encourage the foundation to develop a strategy for tackling this problem head on—foundation-supported schools not only need to recruit many new teachers with both pedagogical and youth development skills but also need to provide a work environment conducive to retaining those teachers over multiple years.

Human resource strategies for new high schools could be developed through a study panel or a targeted grants competition. Potential strategies include splitting responsibilities, as is done at one high school we visited, where counselors handle students' nonacademic issues so that the teachers can focus on academics. Another strategy would be to provide teachers with additional salary, collaboration, and advancement opportunities as both recognition and incentive for doing an extraordinary job.

- **Design and provide a student data system that foundation-supported schools can use to track their students' performance and outcomes.** As noted above, if the foundation wants to measure the initiative's progress with respect to the goal of producing college-ready high school graduates, longitudinal student-level data will be needed. Currently, many of the foundation-supported schools and their districts do not have an effective system for keeping track of individual students and their outcomes after they leave a school. Rather than have each school or district develop its own system, the foundation should fund the design and implementation of a data system that all its funded schools would use for producing regular reports. This activity should include in-service training and technical assistance for school staff on how to use the system, not just for reporting to the foundation but also for monitoring and improving practices to attain better outcomes for all their students.

- **Rethink the school redesign strategy.** Although there have been some isolated examples of apparently successful small schools emerging from the restructuring of a large high school, these have been the exception rather than the rule. On the whole, the data that we have for school redesign efforts are not encouraging. In many cases, the impetus for redesign has come from outside the school, and school staffs have been far from unanimous in their support for the effort. Wrangling over SLC definition and identity, logistics, and the subdivision of students and staff is time-consuming and divisive. Moreover, as noted in Shear et al. (2005), there is some indication that segregation by ability level, language proficiency, and so forth, can be an unintended consequence of the redesign. All of these challenges suggest that the "bar" for funding future school redesign efforts should be set higher than it has been in the past. Evidence of school- and community-level impetus for the redesign, strong leadership from more than one individual, solid curriculum plans and resources, and a realistic plan for dealing with "legacy" students and teachers should be required elements of a redesign plan.

- **Adjust the time frame for school support to match the pace of school establishment and change.** Support for individual high schools under this initiative was limited to 3 years. For a new school starting with a ninth grade and adding a single grade each year, the financial support and technical assistance provided by a foundation grantee ended before the school graduated its first senior class. Every year during which the school received support was a year in which it had a new grade level that it had never before offered. Typically, neither staff nor programs nor curricula had reached a point of stability when the foundation's funding ended. For schools undergoing redesign, both the experience of schools supported through this initiative and prior research suggest that 3 years is insufficient for deep-rooted change (Borman, 2005). We recommend that the foundation consider options for providing schools with 5 years of support, but not automatic support for such an extended term. The foundation could ask its grantees to submit evidence of individual schools' progress after 2 to 3 years of operation, and this evidence could be weighed in determining whether an individual school warrants a fourth and fifth year of support.

- **Continue to analyze and reflect on implementation and student outcome data at the initiative level.** As noted in chapter 1, the foundation's education initiative has evolved in ways that have changed the mix of grants. These changes are leading to support for schools with a broader range of philosophies than was found among the early cohorts of foundation-supported new schools. The more teacher-directed approaches of organizations such as the Knowledge Is Power Program or Chicago Charter School Foundation may produce quite different outcomes than the radically student-centered approaches of the Big Picture Company or EdVisions. Our data suggest that students' academic attitudes are correlated with the foundation's effective-school attributes, as measured by our implementation index, which is a composite of features compatible with the student-centered approaches pursued by the early grantee organizations (Shear et al., 2005). Independent analyses of student and school outcomes over the next 5 years are needed to ascertain whether the more traditional instruction and longer school days used by some of the new grantee organizations will produce the same attitudinal outcomes and whether these schools' academic outcomes will be better than or similar to those of the first set of new schools.

  In addition, some of the more highly specified new school designs promoted by organizations with a national reach are at a point of maturity where a randomized field trial, permitting a rigorous test of the hypothesis that the school model causes better student outcomes, is feasible. Such an experiment would require the foundation's active involvement in recruiting participating districts and structuring school implementation grants to be compatible with the demands of random assignment. An experimental demonstration of a school model's effectiveness would stimulate districts and states around the country to emulate the approach. Thus, such an investment in research would complement the foundation's advocacy efforts. For the school redesign strategy, we have data for no more than 2 years following a school's reorganization. We have found that 2 years is too little time to effect dramatic change in a school's culture, instructional practices, and performance expectations, but we do not know what will happen over a longer term. More could be learned by following the small learning communities through their third and fourth years after redesign.

- **Clarify grantee roles.** As noted above, some of the organizations receiving foundation grants began this work already possessing strong community ties and political influence in a local jurisdiction. Some had experience starting a model school or consulting on school redesign, and some had a particular instructional model and associated resources (AIR/SRI, 2004). Few if any of the grantee organizations possessed all of these capacities, yet fundamentally changing high school education in a locality requires all of them. We encourage the foundation to think about distinct types of grantee organizations and to fund consortia of institutions that collectively offer all the needed expertise (e.g., a strong local partner, a school technical assistance provider, and a partner with national expertise in curriculum and instruction), rather than funding individual organizations. This consortium approach would be likely to entail staged funding, with planning and capacity building in phase-one grants. A phased grant structure would also provide time for building the kinds of partnerships with local institutions of higher education, internship sites, libraries, and museums needed to round out the resources available to small schools with limited budgets. It would also permit more realistic timetables for recruiting school staff, locating school sites, and developing programs and curricula. Finally, the extra time would facilitate the recruitment of local individuals with influence who are willing to champion the new schools and protect them from potentially harmful actions on the part of school boards, district offices, and other organizations.

The U.S. education system as a whole and the American high school especially are notoriously impervious to meaningful change (Hess, 2004; McLaughlin & Talbert, 2001). There was no reason to think that an effort to reinvent the high school would be easy or quickly accomplished. The first 5 years of the Bill & Melinda Gates Foundation's high school grants program have shown that even with a major financial investment, these cautions still hold. Nonetheless, we have seen enough healthy new schools to demonstrate that high school can be a nurturing, inspiring place for historically underserved students. There are also policymakers at the district, state, and national levels who are talking about serious reform of the secondary education system and looking critically at graduation rates for low-income, African-American, and Hispanic/Latino students. Policies that would have been branded utopian just 5 years ago are now receiving endorsements from the nation's governors. Although the work is very far from done, a real beginning has been made.

References

AIR/SRI. (2003). *High time for high school reform: Early findings from the evaluation of the National School District and Network Grants Program.* Washington, DC: Authors. Retrieved from http://www.gatesfoundation.org/nr/downloads/ed/smallschools/Small_schools_eval_2003.pdf

AIR/SRI. (2004). *The National School District and Network Grants Program: Year 2 evaluation report.* Menlo Park, CA: SRI International.

AIR/SRI. (2005). *Evaluation of the Bill & Melinda Gates Foundation's National School District and Network Grants Program. Achievement study: 2004-05. Demographic and student outcome information* [Individual reports for nine districts]. Washington, DC: American Institutes for Research.

Akey, T. M. (2006). *School context, student attitudes and behavior, and academic achievement: An exploratory analysis.* New York: MDRC.

Allensworth, E., & Easton, J. Q. (2005). *The on-track indicator as a predictor of high school graduation.* Chicago: Consortium on Chicago School Research.

Berends, M., Bodilly, S., & Kirby, S. N. (2002). Looking back over a decade of whole-school reform: The experience of New American Schools. *Phi Delta Kappan, 84*(2), 168-175.

Bill & Melinda Gates Foundation. (n.d.-a). *Helping all students achieve.* Seattle, WA: Author.

Bill & Melinda Gates Foundation. (n.d.-b). *High schools for the new millennium: Imagine the possibilities.* Seattle, WA: Author. Retrieved July 20, 2005, from http://www.gatesfoundation.org/nr/downloads/ed/EdWhitePaper.pdf

Bill & Melinda Gates Foundation. (2004). *Creating a portfolio of great high schools.* Seattle, WA: Author. Retrieved July 20, 2005, from http://www.gatesfoundation.org/nr/downloads/ed/GreatHighSchools.pdf

Bill & Melinda Gates Foundation. (2005a). *High-performing school districts: Challenge, support, alignment, and choice.* Seattle, WA: Author.

Bill & Melinda Gates Foundation. (2005b, February). *National Summit on High Schools: Prepared remarks by Bill Gates.* Presented at the 2005 National Education Summit on High Schools, Washington, DC. Retrieved April 7, 2006, from http://www.gatesfoundation.org/MediaCenter/Speeches/BillgSpeeches/BGSpeechNGA-050226.htm

Bill & Melinda Gates Foundation. (2006). *Education program fact sheet.* Seattle, WA: Author. Retrieved February 5, 2006, from http://www.gatesfoundation.org/Education/RelatedInfo/EducationFactSheet-021201.htm

Borman, G. D. (2005). *National efforts to bring reform to scale in high-poverty schools: Outcomes and challenges.* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Bryk, A. S., Nagaoka, J. K., & Newmann, F. M. (2000). *Chicago classroom demands for authentic intellectual work.* Chicago: Consortium on Chicago School Research.

Colby, S., Smith, K., & Shelton, J. (2005). *Expanding the supply of high-quality public schools.* San Francisco: The Bridgespan Group.

Evan, A. M., & Song, M. (2006, April). *Promoting student success at reforming high schools.* Presentation at the annual meeting of the American Educational Research Association, San Francisco.

Foundation Strategy Group. (2005). *Bill & Melinda Gates Foundation: Discussion of preliminary recommendations.* San Francisco: Author.

Hess, G. A. (2004). Districtwide strategies of high school reform: The case of Chicago. In *Using rigorous evidence to improve policy and practice: Colloquium report* (pp. 55-78). New York: MDRC.

Hill, P. T. (2006). Bill and Melinda Gates shift from computers in libraries to reform in high schools. *Education Next, 6*(1), 44-50.

McLaughlin, M. W., & Talbert, J. E. (2001). *Professional communities and the work of high school teaching.* Chicago: University of Chicago Press.

Mirel, J. (2001). *The evolution of the New American Schools: From revolution to mainstream.* Washington, DC: Thomas B. Fordham Foundation.

Mitchell, K., Shkolnik, J., Song, M., Uekawa, K., Murphy, R., Garet, M., & Means, B. (2005). *Rigor, relevance, and results: The quality of teacher assignments and student work in new and conventional high schools.* Washington, DC: American Institutes for Research.

National Center for Education Statistics. (2005). Chapter 2: Elementary and secondary education. In *Digest of education statistics, 2004.* Washington, DC: U.S. Department of Education. Retrieved from http://nces.ed.gov/programs/digest/d04/index.asp

Newmann, F. M. (1992). *Student engagement and achievement in American secondary schools.* New York: Teachers College Press.

Newmann, F., Bryk, A. S., & Nagaoka, J. K. (2001). *Authentic intellectual work and standardized tests: Conflict or coexistence?* Chicago: Consortium on Chicago School Research.

Newmann, F. M., Lopez, G., & Bryk, A. S. (1998). *The quality of intellectual work in Chicago schools: A baseline report.* Chicago: Consortium on Chicago Schools Research.

Parthenon Group. (2002). *Gates Foundation: Assessment of the "tipping point."* Boston: Author.

Provasnik, S., & Dorfman, S. (2005). *Mobility in the teacher workforce: Findings from The Condition of Education 2005* (NCES Publication No. 2005-114). Washington, DC: National Center for Education Statistics.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Rhodes, D., Smerdon, B., Burt, W., Evan, A., Martinez, B., & Means, B. (2005). *Getting to results: Early outcomes in new and redesigned high schools.* Washington, DC: American Institutes for Research.

Shear, L., Song, M., House, A., Martinez, B., Means, B., & Smerdon, B. (2005). *Creating cultures for learning: Supportive relationships in new and redesigned high schools.* Washington, DC: American Institutes for Research.

Vander Ark, T. (2003). America's high school crisis: Policy reforms that will make a difference. *Education Week, 22*(29), 41, 52.

Evaluation of the Bill & Melinda Gates Foundation's High School Grants Initiative: 2001-2005 Final Report

Technical Appendixes

Appendix A. Data Collection and Measurement

Survey Data Collection

The survey data presented in this report come from the principal, teacher, student, and school surveys administered in the springs of 2002, 2003, 2004, and 2005. Model schools were surveyed in the first year of administration only (2002). New schools were surveyed for 3 consecutive years, and comparison schools were surveyed for only 1 year, in either 2004 or 2005. Preredesign schools generally were surveyed first in their preredesign state and then again 2 years later in their redesigned state.

Table A-1 lists the numbers of schools surveyed by school type, year of survey administration, and response rate category. The analytic samples for this report included only schools with adequate response rates. To be considered adequate, both teacher and student response rates had to be at or above 50%, with at least one of the two rates at or above 60%, for all types of schools except comparison schools. The corresponding criteria for comparison schools were that both student and teacher response rates had to be at or above 40%, with at least one of the two rates at or above 50%.

Table A-1. Number of Schools Surveyed, by School Type, Year of Survey Administration, and Response Rate Category

	2002			2003			2004			2005		
	Response Rate Criteria			Response Rate Criteria			Response Rate Criteria			Response Rate Criteria		
School Type	Met	Not Met	Total	Met	Not Met	Total	Met	Not Met	Total	Met	Not Met	Total
Model	5	0	5	0	0	0	0	0	0	0	0	0
New	7	2	9	21	0	21	27	1	28	22	1	23
Comparison	0	0	0	0	0	0	3	2	5	4	1	5
Preredesign	8	1	9	2	1	3	4	1	5	2	1	3
Redesigned	0	3	3	0	0	0	26	1	27	1	0	1
Total	20	6	26	23	1	24	60	5	65*	29	3	32

* The large increase in the number of schools surveyed from 2003 to 2004 was due largely to the fact that each preredesign school in 2002 had broken into multiple redesigned schools by 2004.

Survey Measures

In this section, we describe two groups of measures we derived from the teacher and student survey data: measures of implementation and student attitudes toward schooling.

Implementation Index. We created an implementation index to indicate the overall level of implementation of effective-school attributes. Specifically, we first used factor analyses to construct a set of scales that were mapped onto 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.[49] The relevant survey items comprising each of the scales and the reliabilities of the scales in each survey year are listed in Table A-2. On the basis of these scales, we created measures of the six attributes of effective schools and the overall implementation index as follows:

- Aggregate the teacher scales and student scales 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 among all schools with adequate survey response rates.
- Create the implementation index as the mean of the six standardized attribute measures.
- Standardize the implementation index among all schools with adequate survey response rates.

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

Table A-2. Scales Measuring Effective-School Attributes, Survey Items Used in the Scales, and Alpha Coefficients

Scale	School Attributes and Survey Items	Reliability (α) 2002-05
Common Focus		
Common Focus (Teacher)	♦ How much agree: Teachers have different visions for student learning ♦ How much agree: Teachers share beliefs about what the central mission of the school should be ♦ How much agree: Teachers are committed to developing strong relationships with students ♦ How much agree: Teachers committed to developing partnerships with parent(s)/guardian(s) for student learning ♦ How much agree: Parent and community members do share vision for student learning	.77 - .83
Instructional Coherence (Teacher)	♦ How much agree: Support programs linked curricula, instruction, and assessments ♦ How much agree: Professional development supports the implementation of a set of common curricula, instructional strategies, and assessments ♦ How much agree: Curricula coordinated to avoid repeating subject matter with students as they more from grade to grade ♦ How much agree: Familiar with curricula and instructional strategies used by colleagues who are also teaching my students in subject areas other than my own ♦ How much agree: Teachers have adequate opportunity to meet with one another	.67 - .81

[49] 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.

Table A-2. Scales Measuring Effective-School Attributes, Survey Items Used in the Scales, and Alpha Coefficients (*continued*)

Scale	School Attributes and Survey Items	Reliability (α) 2002-05
High Expectations		
High Expectations (Teacher)	♦ How much agree: Most teachers: Set high standards for teaching ♦ How much agree: Most teachers: Set high standards for students learning ♦ How much agree: Most teachers: Make expectations for instructional goals clear to students ♦ How much agree: Most teachers: Carefully track students' academic progress	.88 - .89
High Expectations (Student)	♦ How much agree: Teachers at school: Believe all students can do well ♦ How much agree: Teachers at school: Given up on some students ♦ How much agree: Teachers at school: Care only about smart students ♦ How much agree: Teachers at school: Expect very little from students ♦ How much agree: Teachers at school: Work hard to make sure all students are learning	.71 - .76
Personalization		
Personalization—Social (Teacher)	♦ Percentage of students for whom know: Their first and last names ♦ Percentage of students for whom know: Their academic aspirations ♦ Percentage of students for whom know: Their academic background prior to this year ♦ Percentage of students for whom know: Their home life ♦ Percentage of students for whom know: Names of person/people with whom they live ♦ Percentage of students for whom know: Who their friends are ♦ Percentage of students for whom know: Their cultural and linguistic background	.92 - .93
Personalization—Academic (Teacher)	♦ Extent to which help students with academic difficulties by: Diagnosing problems the students are having ♦ Extent to which help students with academic difficulties by: Determining how to match school resources to student needs ♦ Extent to which help students with academic difficulties by: Gathering info to help understand students' difficulties ♦ Extent to which help students with academic difficulties by: Helping student learn how to overcome their difficulties in ways that compensate for different learning disabilities	.88 - .89

Table A-2. Scales Measuring Effective-School Attributes, Survey Items Used in the Scales, and Alpha Coefficients *(continued)*

Scale	School Attributes and Survey Items	Reliability (α) 2002-05
Personalization		
Personalization—School Action (Teacher)	♦ Extent to which your school provides following help to students with academic difficulties: Extra attention from you ♦ Extent to which your school provides following help to students with academic difficulties: Extra help from other staff member during regular school day, week, or year ♦ Extent to which your school provides following help to students with academic difficulties: Extra help from school staff outside regular school day, week, or year ♦ Extent to which your school provides following help to students with academic difficulties: Parent-teacher meetings to discuss what the school and the student's parent(s)/guardian(s) can do to help ♦ Extent to which your school provides following help to students with academic difficulties: Referrals to community organizations for assistance ♦ Extent to which your school provides: Extra help from other students	.81 - .85
Personalization (Student)	♦ How many adults in your school: Willing to give extra help with your homework if needed ♦ How many adults in your school: Willing to help you with a personal problem ♦ How many adults in your school: Really care about how well you are doing in school ♦ How many adults in your school: Have helped you think about whether you are meeting the requirements for graduation ♦ All teachers/adults willing: Help you think about what you need to do to prepare for college or a career	.84 - .86
Collaboration		
Time to Collaborate (Teacher)	♦ How often have you engaged in: Observing other teachers while they teach ♦ How often have you engaged in: Being observed by other teachers while teaching ♦ How often have you engaged in: Receiving feedback from other teachers based on their observations of your teaching ♦ How often have you engaged in: Providing feedback to other teachers based on your observations of their teaching ♦ How often have you engaged in: Coteaching or mentoring other teachers or staff in your school ♦ How often have you engaged in: Diagnosing individual students' learning with other teachers	.80 - .82

Table A-2. Scales Measuring Effective-School Attributes, Survey Items Used in the Scales, and Alpha Coefficients *(continued)*

Scale	School Attributes and Survey Items	Reliability (α) 2002-05
Collaboration		
Reflective Professional Dialogue (Teacher)	♦ How often met with other teachers to discuss: The goals of this school ♦ How often met with other teachers to discuss: The structure of the school day ♦ How often met with other teachers to discuss: Development of new curricula or modification of existing curricula ♦ How often met with other teachers to discuss: Teaching practices or instructional issues ♦ How often met with other teachers to discuss: General classroom administration and management	.86 - .87
Parent Involvement (Teacher)	♦ How often have you: Involved parent/guardian in setting up particular learning objectives for student ♦ How often have you: Involved parent/guardian in judging student Work ♦ How often have you: Provided parent/guardian with exemplars of excellent student work to demonstrate standards for good performance ♦ How often have you: Involved parent/guardian as mentors for individual students or groups of students	.79 -.85
Community Resources (Teacher)	♦ 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 ♦ How often in your selected instructional period: Discussed different cultures in your community ♦ How often in your selected instructional period: Taken students to visit places or organizations in the community	.57 - .62
Respect and Responsibility		
Respect and Responsibility (Teacher)	♦ How much agree: Teachers feel good about parents/guardians' support of their work ♦ How much agree: Students treat one another with respect ♦ How much agree: Relationships between students and teachers based on mutual trust and respect ♦ How much agree: Students get teased if they take academics seriously ♦ How much agree: Student success/failure is due to factors beyond teachers' control ♦ How much agree: I can usually get through to even the most difficult students ♦ How much agree: It is the responsibility of teachers to keep students from dropping out ♦ How much agree: Teaching makes a difference in students' lives	.71 - .75

Table A-2. Scales Measuring Effective-School Attributes, Survey Items Used in the Scales, and Alpha Coefficients *(continued)*

Scale	School Attributes and Survey Items	Reliability (α) 2002-05
Respect and Responsibility		
Respect and Responsibility (Student)	♦ How much agree: Many students in this school do not respect one another ♦ How much agree: There are groups of students in this school who do not get along ♦ How many students: Feel OK to make racist or sexist remarks ♦ How many students: Feel OK to cheat ♦ How many students: Feel OK to get into physical fights ♦ How many students: Feel OK to steal things from students ♦ How many students: Feel OK to destroy or steal school property	.82 - .84
Collegiality (Teacher)	♦ How much agree: Teachers really don't support each other or work together ♦ How much agree: Teachers in this school trust and respect one another ♦ How much agree: Teachers, administrators, and other staff at this school model responsible behavior for students to see	.73 - .79
School Climate Safe (Teacher)	♦ How often have you felt unsafe: In your classes ♦ How often have you felt unsafe: In hallways, stairs, and bathrooms ♦ How often have you felt unsafe: Immediately outside the school	.87 - .94
School Climate Safe (Student)	♦ How often have you felt unsafe: In your classes ♦ How often have you felt unsafe: In hallways, stairs, and bathrooms ♦ How often have you felt unsafe: Immediately outside the school	.85 - .88
Orderly Climate (Student)	♦ How often has this occurred in school: Fighting ♦ How often has this occurred in school: Destroying property ♦ How often has this occurred in school: Verbal bullying ♦ How often has this occurred in school: Physical bullying ♦ How often has this occurred in school: Cheating ♦ How often has this occurred in school: Theft	.91 - .93
Technology as a Tool (Teacher)	♦ How often do your students use technology for: Expressing themselves in writing ♦ How often do your students use technology for: Communicating electronically about academic subjects ♦ How often do your students use technology for: Exploring ideas and information ♦ How often do your students use technology for: Analyzing information ♦ How often do your students use technology for: Presenting information to an audience ♦ How often do your students use technology for: Improving computer skills	.89 - .90

Student Attitudes Toward Schooling. On the basis of student surveys, we created the following measures of student attitudes:

- **Educational engagement:** a measure of how often students asked questions in class or contributed to class discussions, met with teachers to talk about schoolwork, talked to family about schoolwork, asked friends for advice about schoolwork, and worked with classmates outside of class or school on schoolwork
- **Persistence:** a measure of how often students made extra effort on challenging assignments, got help with difficult homework, and tried to do well on schoolwork even if it was too hard or not interesting
- **Academic self-concept:** a measure of the degree to which students felt they were good at reading, writing, learning mathematics, getting help, and working with others
- **Educational aspirations:** a measure of how far students plan to go in schooling, ranging from dropping out of high school to post-college education

Higher values on each of these indices indicate more positive student attitudes. The specific survey items comprising these indices are listed in Table A-3.

Table A-3. Student Attitude Scales, Survey Items Used in the Scales, and Alpha Coefficients

Scale	Survey Items	Reliability (α) 2002-05
Engagement—Interest (Student)	♦ This school year, I have talked to my family about what I am working on in school. ♦ This school year, I have asked my friends for advice about something I am working on in school. ♦ This school year, I have asked questions in class or contributed to class discussions. ♦ This school year, I have worked with classmates outside of class or school on schoolwork. ♦ This school year, I have asked my teachers to meet with me to talk about grades, assignments, or my work on projects.	.75 - .76
Engagement—Persistence (Student)	♦ 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.	.74 - .76

Table A-3. Student Attitude Scales, Survey Items Used in the Scales, and Alpha Coefficients *(continued)*

Scale	Survey Items	Reliability (α) 2002-05
Academic Self-Concept (Student)	♦ I am good at asking teachers for help when I get stuck on schoolwork. ♦ I am good at working in a group with other students. ♦ I am good at taking part in class or group discussions. ♦ I am good at understanding what I read. ♦ I am good at writing papers or stories. ♦ I am good at learning mathematics.	.68 - .70
Educational Aspirations (Student)	♦ Right now, what is your best guess about how far you will go in school?	N/A

Site Visit Data Collection

Site Visits

In 2004-05, site visits were conducted in 35 of the schools in the survey sample. Twenty-three newly opened schools, four schools prior to redesign, and eight redesigned schools were visited. Table A-4 shows the types and numbers of schools that were visited in 2004-05, by the number of years of implementation.

Table A-4. Number of Site Visit Schools, by School Type and Implementation Year

Type of School	Fourth Year	Third Year	Second Year	First Year	Total
New	5	7	6	5	23
Preredesign	N/A	N/A	N/A	4	4
Redesigned	5	0	3	N/A	8
Total	10	7	9	9	35

Two-person teams visited each school over a period of 2 to 4 days. At redesigned schools, site visitors collected data at two small learning communities (SLCs). 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 newly opened schools, redesigned schools, and schools prior to redesign. Interviews and focus groups were audiotaped to support the completeness and accuracy of the data records.

Principals and Lead Staff Interviews

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.

Teacher Interviews

Site visit teams interviewed five teachers at each school. Teachers to be interviewed were selected according to the following criteria:

- 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 (ELA) teacher (if the school didn't have a 10th-grade ELA teacher, we interviewed a 9th- or 11th-grade ELA 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 ELA or mathematics classes; in these cases, we asked leaders to identify teachers of classes where mathematics and ELA 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). The teacher interview protocols covered topics such as relationships among teachers and between teachers and students, the school's learning environment, and the school's ability to serve all students well.

Student Focus Groups

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 students 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 was 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 was preparing them for life after graduation.

Classroom Observations

The site visit teams conducted 50-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 briefly with teachers to discuss what they had seen, typically asking 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.

Building and Structure Inspections

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.

Site Visit 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 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. Experienced 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. Code categories included key school attributes, school capacity issues, school organization and climate, curriculum and classroom practice, and outcomes for students and schools. 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 159 codes in all.

Data coding began with test coding, moved on to reliability and validity coding, and concluded with operational coding. Twelve coders were trained to use the new 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 a senior analyst 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 two to four 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 three senior analysts, who then developed a set of master codes for the main ideas of the paragraph; these codes were negotiated with the coding team. The resulting set of codes 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 the agreement rate was calculated as follows: agreements/(agreements + disagreements). The reasons for any low agreement 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 weekly intervals to develop our final sample for reliability and validity. In the cumulative sample from six reliability runs and more than 4,000 individual applications of the codes to data, the 12 coders had an average rate of 81% agreement with the standard codes established by the senior analysts. Among the codes, 82% of codes that were used more than four 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, which often included discussions of the school mission and goals. Such interrelated constructs made coding distinctions challenging. In cases like these, we computed reliability estimates for two interrelated codes together and asked analysts to query 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.

Site Visit Data Analysis

During operational coding, coders entered their coding work into a database using a qualitative analysis software program called ATLAS.ti. Many of the analysts began their work by reviewing information about the school, including reviewing the school summary written by site visitors; a brief descriptive table on school type, size, and context; and/or samples of data capture forms for schools in their analysis group. These reviews helped analysts develop a more comprehensive view of the school contexts and schoolwide issues.

Analysts then queried the database to review coded data by topic. In some cases, they used coded data to find examples of issues that surfaced in analyses of survey data. 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 examined the coded data on a given 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.

Demographic, Attendance, Progression Rate, and Student Achievement Data

For the analyses reported in chapter 2, we collected extant data on demographic characteristics (free or reduced-price lunch status; minority status; English language learner status; and students requiring special educational services), attendance rates, progression rates, and achievement scores from foundation-supported schools and other public high schools in the same district. For the analyses, we used data from the districts shown in Table A-5.

Table A-5. Data Availability for Chapter 2 Analyses, by District

District	Demographics	Attendance	Progression	Achievement Scores	Proficiency Rates
Anderson Union, CA				X	
Baltimore, MD[50]	X	X	X	X	
Chicago, IL	X	X	X	X	X
Cincinnati, OH	X	X	X	X	
Cleveland, OH				X	
Cotati-Rohnert Park, CA				X	
El Dorado, CA				X	
Los Angeles, CA	X				
New York City, NY[51]	X	X		X	
Oakland, CA	X	X		X	X
Providence, RI	X	X	X	X	X
Ravenswood, CA					
Sacramento, CA	X			X	X
San Diego, CA	X				

These districts vary in the scope and depth of their respective reform efforts and in the type and number of schools within them. Individual school districts also vary with regard to the types of student assessment instruments they employ, their data collection methods, their respective definitions of data elements, and the political contexts surrounding district decision-making.[52] Given these differences, we have generally restricted our comparisons to be among schools within the same district.

Depending on the type of reform attempted, foundation-supported schools were categorized into two school types:[53]

- **New—**Schools that were newly created using foundation support at least in part.
- **Redesign—**Schools that received foundation support for redesign and had redesigned into smaller schools.

[50] Although officially called "Baltimore City," we refer to this district as "Baltimore" in this report.
[51] Three (Brooklyn, the Bronx, and Manhattan) of the five New York City boroughs were examined. These were the three boroughs with foundation-supported schools.
[52] Districts also varied in the degree to which they were able to provide student achievement data.
[53] There was a third type of foundation-supported schools called preredesign. However, because preredesign schools had not yet begun the process of conversion, they were not included in the analyses.

The types of high schools within the districts named above varied considerably. All of the districts had at least one new school, and all but one district (New York City) had both new schools and redesigned schools.[54] The analyses were focused on new vs. the district average (an average of school averages) and redesigned vs. the district average. For the analysis of student achievement scores, we collected data from the same districts discussed above.

We had two measures of student high school achievement: one for ELA and one for mathematics, from the 2003-04 school year. To assess the relationship between student achievement and foundation support, we reviewed differences in mean scores on achievement tests and determined whether students at each of the schools performed above their district's mean ELA and math assessments. We restricted the analysis to those schools in which we could 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 school's assessment results in proper context. State assessments have different metrics and are mapped to state-specific standards. Therefore, direct comparisons of students in different states are inappropriate. Instead, to highlight students' relative success, we compared each school's mean student test scores with the performance of students in the district as a whole. We limited our analysis to sample schools for which we could obtain the district's mean scores.

To conduct the analyses, 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 or below district mean) and tallied the number of new and redesigned schools that fell above the district mean on each of these achievement measures and calculated a percentage above the district mean based on these tallies.

Teacher Assignment and Student Work Data

To measure students' opportunities to learn, we collected eight assignments that teachers gave to students in class or as homework over the course of the school year. We asked for four assignments that were typical of day-to-day work and four that were particularly challenging and offered opportunities for students to produce high-level work. We randomly selected 10 students from the eligible 10th-grade students in the class and collected the work they did in response to three of the eight assignments (Mitchell et al., 2005). For each assignment, we measured the level of rigor and the level of real-world relevance. We also measured the overall quality of each piece of student work.

We collected these data from 10th-grade ELA and mathematics teachers from three types of schools over 3 school years. The graphs in the report present data collected from 12 new small schools from across the United States for 2 years—2003-04 and 2004-05—

[54] Baltimore and New York City had high concentrations of selective/specialty schools among their non-foundation-supported schools, with 10 out of 27 high schools in Baltimore and 123 out of 201 in NYC. Given these high concentrations, we paid additional attention to comparisons involving these types of schools in the individual achievement study reports for these two cities.

and from 8 large comparison schools that are geographically and demographically similar to the 12 new schools. The analyses also include data from 11 preredesign schools, although they are not reported in the text. These 11 schools increased our sample size, but data were collected for analysis in our next report, which will examine changes in rigor, relevance, and student work quality after schools were redesigned into small learning communities.[55] The comparison school assignments and work were collected in 2004-05. Assignments and work were collected from the national preredesign schools in 2003-04, and data for the eight Washington preredesign schools were collected in 2002-03. Sample sizes for the three school types are listed in Table A-6.

Table A-6. Number of Observations, by School Type

	New	Comparison	Preredesign		New	Comparison	Preredesign
Schools	12	8	11	Schools	12	8	11
ELA Assignments				**ELA Student Work**			
Assignments	331	127	228	Student Work	614	284	480
Teachers	38	19	33	Students	346	155	238
Math Assignments				**Math Student Work**			
Assignments	232	135	238	Student Work	494	293	533
Teachers	33	19	31	Students	278	146	240

Scoring of Teacher Assignments and Student Work and the Many-Facet Rasch Model

After data collection was completed each school year, we hired local 10th-grade English/language arts and mathematics teachers to score the assignments and work using rubrics originally conceived by Newmann and Bryk (Newmann et al., 1998; Bryk et al., 2000; Newmann et al., 2001). There were 12 master teachers in each subject for the first two summers, and 19 in the third year (18 in mathematics). Many scorers participated in more than one of the three scoring sessions.

To test for interrater reliability, we picked a subset of assignments and work to be scored twice, each time by a different rater. We examined interrater reliability by calculating how often both raters gave the same score (i.e., perfect agreement) or scores within 1 point of each other when they scored the same criterion of the same assignment or piece of work. Table A-7 lists the agreement rates.

Agreement rates varied by criterion. The ELA scorers had at least 63% perfect agreement on the four assignment criteria and at least 61% on the three student work criteria. They had at least 91% agreement within one point on the four assignment criteria and at least 93% agreement within one point on the three student work criteria. We observed

[55] We collected data from 12 large high schools prior to their redesign into small learning communities. One of these schools was the only remaining large school in the same geographic area as one of the new schools in the sample and is considered a comparison school for the purpose of this report.

slightly more variation in mathematics scores. The mathematics scorers had at least 55% perfect agreement on the four assignment criteria and at least 59% on the three student work criteria. They had at least 85% agreement within one point on the four assignment criteria and at least 79% agreement within one point on the three student work criteria.

Analysis of interrater agreement is important as we continue to refine our training procedures for the scoring of assignments and work, but it is not an analytical concern because we use Many-Facet Rasch Measurement (MFRM), which allows us to account for rater severity when creating our measures.

Table A-7. Agreement Rates on Assignments and Student Work Scored by Two Raters

Criterion	Possible Scores	Perfect Agreement	Agreement Within 1 Point
ELA			
Rigor of teacher assignment			
Elaborated communication	1-4	65%	95%
Construction of knowledge	1-4	63%	91%
Relevance of teacher assignment			
Real-world connections	1-4	82%	91%
Student involvement in crafting assignments	1-4	85%	99%
Student work quality			
Construction of knowledge	1-4	61%	94%
Elaborated communication	1-4	64%	93%
Language conventions and resources	1-6	69%	93%
Mathematics			
Rigor of teacher assignment			
Important mathematical content	1-4	55%	85%
Problem solving and reasoning	1-4	64%	93%
Effective communication	1-3	81%	100%
Relevance of teacher assignment			
Real-world connections	1-4	85%	99%
Student involvement in crafting assignments	1-4	92%	99%
Student work quality			
Conceptual understanding	1-4	59%	85%
Procedural knowledge	1-4	60%	79%
Problem solving and reasoning	1-4	74%	95%
Effective communication	1-4	73%	98%

The product of the scoring session is a data set in which each teacher assignment or piece of student work receives at least one score for each criterion. To create measures that capture the rigor and relevance of teacher assignments and the quality of student work, we used MFRM. As information resources, MFRM uses the differences of the scores given by evaluators and differences in criteria difficulty and evaluators' severity. Doing so differentiates the rigor and relevance of teacher assignments and quality of pieces of student work.

MFRM is based on an ordered logistic regression model; the ratings evaluators used are ordered categories. The following MFRM model predicts an evaluator's rating when he or she is judging the rigor or relevance of a teacher assignment or the quality of a piece of student work:

$$\log\left(\frac{P_{nijk}}{P_{nij(k-1)}}\right) = B_n - C_i - D_j - F_{ik}$$

where

- P_{nijk} is the probability of assignment (or student work) n getting a score of k on criterion i by scorer j
- $P_{nij(k-1)}$ is the probability of assignment (or student work) n getting a score of k-1 on criterion i by scorer j
- B_n is the parameter for the quality of assignment (or student work) n
- C_i is the parameter for rubric i (stringency of the criterion)
- D_j is the parameter for scorer j (severity of the scorer)
- F_{ik} is the parameter for receiving a score of k relative to k-1 on criterion i (step difficulty).

Of the estimates generated by the above model, the most important is B_n.[56] It represents the measures of assignment rigor, assignment relevance, or student work quality. Two criteria comprise the rigor measure in ELA: elaborated communication and construction of knowledge. For mathematics, the measure is based on three criteria: effective communication, problem solving and reasoning, and important mathematical content. Two criteria comprise the relevance measure for ELA: student involvement and real-world connections. The same two criteria are used to measure relevance in mathematics assignments. To derive measures of student work quality in ELA, we used three criteria: language convention and resources, elaborated communication, and construction of knowledge. For mathematics, we used four criteria: procedural knowledge, conceptual understanding, effective communication, and problem solving and reasoning. These measures are presented in Table A-7.

Diagnostics of MFRM Scores

The MFRM model produces linear and objective measures only when certain assumptions are met. In the context of our study, some assumptions seem less tenable than others. Because each year we replaced some of the scorers with new scorers, it was possible that the severity of scoring changed from 1 year to the next. To test this possibility, we selected a subset of teacher assignments and pieces of student work to be evaluated

[56] The other estimates are also important, because they serve as statistical controls; the MFRM corrects the estimates of rigor, relevance, and student work quality for scorer severity and criterion difficulty.

every year. This allowed us to examine whether evaluators varied in terms of their rating severity across years.

The greatest drift in severity of teacher assignment ratings occurred between years 1 and 2 of scoring. There were no significant differences between mean scores on each rubric between years 2 and 3. Student work scores also tended to drift more between years 1 and 2 than years 2 and 3, although we did observe increased leniency in the scoring of multiple criteria in mathematics student work between years 2 and 3. MFRM lessens the impact of this problem by adjusting for rater severity. Raters are treated as separate individuals from year to year. By controlling for rater severity, we also control for any yearly changes in severity.

The MFRM model produces more reliable and accurate estimates when there is a spread of difficulty among criteria. A larger variance in criterion difficulties leads to more robust MFRM scores. Table A-8 presents difficulty estimates for the evaluation criteria.

Table A-8. Ranking of Criterion Difficulty

Criterion	Measure	Criterion Difficulty
ELA		
Teacher assignments		
Elaborated communication	Rigor	-1.84
Construction of knowledge	Rigor	-0.33
Student involvement	Relevance	1.00
Real-world connections	Relevance	1.16
Student work		
Language conventions and resources	Student work quality	-0.24
Elaborated communication	Student work quality	0.01
Construction of knowledge	Student work quality	0.23
Mathematics		
Teacher assignments		
Problem solving and reasoning	Rigor	-0.37
Important mathematics content	Rigor	-0.25
Effective communication	Rigor	-0.23
Real-world connections	Relevance	0.28
Student involvement	Relevance	0.58
Student work		
Conceptual understanding	Student work quality	-0.79
Procedural knowledge	Student work quality	-0.58
Effective communication	Student work quality	0.56
Problem solving and reasoning	Student work quality	0.82

As seen above, item difficulty estimates for ELA assignments were more widely distributed than those for mathematics assignments. Although lower variation in items does decrease the sensitivity of the MFRM model, it is important to note that our MFRM also uses other sources of information: rater severity differences and difficulty differences in evaluators' responses. Thus, the impact of a limited score spread within particular rubrics on our MFRM estimates is lessened.

Table A-9 presents reliability measures for rigor, relevance, and student work quality in ELA and mathematics. These reliability measures are similar to Cronbach's alpha.

Table A-9. Reliability of Rigor, Relevance, and Student Work Quality Estimates

	Reliability (α)
ELA	
Assignment rigor	.69
Assignment relevance	.35
Student work quality	.72
Mathematics	
Assignment rigor	.64
Assignment relevance	.60
Student work quality	.43

Researchers generally set a threshold for acceptable reliability at .65 to .70. Because the reliability approached but did not always meet this threshold, we used latent variable hierarchical linear modeling (HLM). HLM decreases the influence of unreliable scores by giving them lower weights in the analysis. Chapter-specific appendixes below describe the analyses in more detail.

Creating MFRM Score Categories

To allow intuitive interpretation of the MFRM scores, we divided the three outcomes of interest—assignment rigor, assignment relevance, and student work quality—into four categories: little to none, limited, moderate, and substantial.

To categorize each assignment and piece of work into one of these four levels, it was necessary to determine how raw scores from each criterion would correspond to each category of teacher assignment or student work quality. For most of the criteria, the translation was relatively simple; most criteria were scored on a four-point scale. In these cases, a score of 1 would correspond to "little to none," a score of 2 would correspond to "limited," and so on. For the criteria that were scored on a three- or six-point scale, we consulted the scoring leaders, who helped us determine how these raw scores would be translated. For example, the mathematics scoring leader indicated that for criterion 3, which was scored on a three-point scale, a score of 1 would indicate little to no rigor, a 2 would indicate limited or moderate rigor, and a 3 would indicate substantial rigor.

The next step in the translation of assignment and work scores to quality categories was the calculation of average MFRM scores corresponding to each raw score obtained for each criterion. Using mathematics rigor as an example, we report these averages in Table A-10. The average MFRM score for a raw score of 1 was 3.12 for criterion 1, 3.24 for criterion 2, and 3.47 for criterion 3. From these average values corresponding to a raw score of 1 on each criterion, we obtained an average MFRM score across the three criteria: 3.28. Using the same procedure, we derived 4.68, 5.68, and 7.06 as the average MFRM values corresponding to raw scores of 2, 3, and 4.

Table A-10. Translating MFRM Scores for Rigor Mathematics Rigor into Four Categories

Criteria	Average Raw Score for Each Category			
	1. Little to No	2. Limited	3. Moderate	4. Substantial
1. Important math content	3.12	4.26	5.93	7.07
2. Problem solving and reasoning	3.24	4.67	6.00	7.45
3. Effective communication	3.47	5.12	5.12	6.67
Average score	3.28	4.68	5.68	7.06

We used the midpoint values between the average scores as the cutoffs between categories. In the case of mathematics rigor, as shown in Table A-11, the threshold values are 3.98 (the midpoint between the average MFRM scores for categories 1 and 2), 5.18 (the midpoint between the average MFRM scores for categories 2 and 3), and 6.37 (the midpoint between the average MFRM scores for categories 3 and 4).

Table A-11. Cutoffs for the Four Categories: Mathematics Rigor

Category	Cutoff
Little to no	0 to < 3.98
Limited	3.98 to < 5.18
Moderate	5.18 to < 6.37
Substantial	6.37 to 10.00

Student Achievement Score Data Linked to Student Work

To assess the impact of student work on student achievement, we collected demographic data and 9th- and 10th-grade achievement scores in ELA and mathematics from participating students. Because students at different schools took tests that were often not comparable, we converted the scores into national normal curve equivalent (NCE) scores where possible. The national tests for which norming information was available were the CAT-6 (California) and the SAT-9 (Minnesota and Rhode Island). By using the national norms to convert the test scores to a common metric, we assumed that different tests captured the same construct and the norming samples were from populations that have the same distribution of ELA or mathematics achievement. The rest of the schools were either unable to provide test score data or were able to provide data only for tests that were not nationally normed and therefore not able to be converted into a common metric with the data from other schools.

We used the 9th-grade test scores as a measure of students' prior achievement. Missing 9th-grade test scores were imputed by using a best-subsets regression model within the Stata statistical software program. Missing data were also an issue for 10th-grade test scores; however we did not impute the dependent variable—10th-grade score.

Appendix B. Analysis for Chapter 2: Promoting Student Success

In this chapter, we compare the performances of foundation-supported schools with those of their host districts. These comparisons allowed us to gauge the performance of the foundation-supported schools along various measures without making comparisons across districts, because districts employ different definitions and different measures of indicators such as attendance and proficiency, comparisons across districts are not particularly meaningful.

The Quality of Student Work in New Schools and Comparison Schools

The focus of this report is new schools and comparison schools. Comparison schools are similar to the new schools geographically and demographically and are where students likely would have enrolled if the new school had not opened. We also included data from preredesign schools in our analyses to increase the precision of the parameter estimates. The analytic sample for the cross-school-type comparisons of student work quality included student work from:

- Twelve new schools
- Eight comparison schools
- Eleven preredesign schools

The final analytic sample for ELA consisted of 1,378 pieces of student work (from 682 students), and the sample for mathematics consisted of 1,246 pieces of student work (from 619 students).

Our data on student work are of a nested nature—i.e., pieces of student work are nested within students—therefore, we used hierarchical liner modeling (HLM) as our primary analytic method, which has been specifically designed for analyzing nested data (Raudenbush & Bryk, 2002). In addition, because the quality of student work is not directly observable and hence inevitably measured with error in our Many-Facet Rasch Measurement (MFRM) analyses, as detailed in the preceding section, we treat the quality of student work as a latent variable and take advantage of the fact that MFRM provides both a measure of student work quality and a measure of the standard error. We use a three-level HLM latent variable model, where level 1 is the measurement model, level 2 is the student work level, and level 3 is the student level.

The purpose of the measurement model at level 1 is to explicitly take into account the measurement errors in the MFRM scores of student work quality in our analyses. We also built in controls for assignment type (challenging vs. typical), student characteristics (i.e., prior achievement, race, and gender), and a school risk index based on percentage minority and percentage of students eligible for free or reduced-price lunch. Separate analyses were conducted for ELA and mathematics.

More specifically, level 1 of the three-level HLM latent variable model is a measurement model, which conceives the MFRM rating of each piece of student work as comprising the true score of the underlying latent measure of student work quality and a measurement error. The true scores of the latent variable, student work quality, estimated at level 1 were then used as the outcomes at level 2 (student work level) and modeled as a function of assignment type. The level-2 intercepts, which represent the average level of work quality for each student, adjusted for assignment type, were further modeled as random effects in the student-level model at level 3. Ideally, we would further nest students within schools. However, HLM software can only accommodate up to three levels, so we added school-level measures to the student-level model at level 3. The specification of the three-level model is as follows:

Level-1 Model (Measurement):

$$Y_{jk} = \pi_{jk} (DUMSW\_WT) + \varepsilon_{jk}$$

where

- The outcome Y_{jk} is the observed MFRM score of the quality of student work j submitted by student k. The MFRM scores were weighted by the inverse of the standard errors of the measurement derived from the MFRM model.
- DUMSW_WT is a dummy variable representing the latent variable that generated the observed scores of student work quality. It was uncentered at level 1 and also weighted by the inverse of the standard errors of the MFRM scores.
- π_{jk} is the true score of the latent variable for student work j submitted by student k.
- ε_{jk} is a measurement error embedded in the observed score; $\varepsilon_{jk} \sim N(0,1)$, given the weighting of both the dependent and independent variables.

Level-2 Model (Student Work):

$$\pi_{jk} = \beta_{0k} + \beta_{1k} (CHALLENGING) + r_{jk}$$

where

- CHALLENGING is a dummy variable indicating the type of teacher assignment that is associated with student work j from student k (1 = challenging, 0 = typical). It was grand-mean centered and fixed at level 3.
- β_{0k} is the mean student work score for student k adjusted for the type of assignment.
- r_{jk} is a random effect associated with student work j submitted by student k on the quality of the work.

Level-3 Model (Students):

$$\beta_{0k} = \gamma_{00} + \gamma_{01}\text{(PRIORACH)} + \gamma_{02}\text{(BLACK)} + \gamma_{03}\text{(ASIAN)} + \gamma_{04}\text{(HISPANIC)} + \gamma_{05}\text{(INDIAN)} + \gamma_{06}\text{(MALE)} + \gamma_{07}\text{(ZRISK)} + \gamma_{08}\text{(NEW)} + \gamma_{09}\text{(PREREDESIGN)} + u_{0k}$$

$$\beta_{1k} = \gamma_{10}$$

where

- PRIORACH is a grand-mean-centered measure of students' ninth-grade achievement. Since students in different schools took different types of standardized tests in ninth-grade, we converted their ninth-grade scores to NCEs based on the norming tables provided by test developers. We also imputed missing data on this measure via best-subsets regression, using student-level eighth-grade test scores, ethnicity, gender, and socioeconomic predictors.[57]
- BLACK, ASIAN, HISPANIC, and INDIAN are dummy variables for student race, with WHITE being the reference group. MALE is a dummy variable for student gender. All these student-level control variables were grand-mean centered.
- NEW, PREREDESIGN, and ZRISK are measures of school characteristics. All students in the same school share the same value on these two measures. ZRISK is a school risk index based on the percentage of minority students and the percentage of students eligible for free or reduced-price lunch. It was centered around its grand mean. NEW and PREREDESIGN are uncentered dummy variables representing new schools and preredesign schools, respectively, with comparison schools being the omitted reference group.
- γ_{00} is the grand mean of the quality of student work across all students, adjusted for assignment type.
- γ_{10} is the average effect of assignment type (i.e., challenging vs. typical) on the quality of student work across all students.
- γ_{08} and γ_{07} represent the school type effects, which are the differences in the quality of student work between small and comparison schools and between preredesign and comparison schools, respectively, adjusted for assignment type, student characteristics, and the school risk index.
- u_{0k} is a random effect associated with student k on the quality of student work.

Tables B-1 and B-2 report the results of the student work quality models. Model 1 is a baseline model, and Model 2 tests whether the relationships found for student work quality and student achievement hold up with a set of statistical controls.

[57] Although the norming samples for different standardized tests are not identical, we assume that they are similar enough to provide a consistent measure for prior achievement as a student-level control variable in the HLM model.

Table B-1. Student Work Quality in ELA, Differences by School Type—HLM Estimates of Fixed Effects and Variance Components

Measure	Model 1 (Baseline)			Model 2 (Full)		
	Coefficient	**t-stat**	**Signif.**	**Coefficient**	**t-stat**	**Signif.**
Intercept	3.60	0.11	***	3.08	12.30	***
New (vs. comparison)	**0.19**	**0.13**		**0.27**	**2.20**	*
Preredesign (vs. comparison)	-0.23	0.13		-0.67	-5.01	***
Challenging	0.93	0.08	***	0.93	11.52	***
School risk index				-0.58	-9.03	***
% lunch assistance				0.15	1.15	
9th-grade test score				0.01	2.84	**
Male				-0.14	-1.53	
Black				0.25	1.60	
Asian				0.15	0.86	
Hispanic/Latino				0.00	0.01	
Indian				-0.15	-0.47	
Variance components:						
Between student assignments	1.01	63%		1.04	74%	
Between students	0.59	37%		0.36	26%	
Total	1.60	100%			100%	

Note. N of schools = 31 (12 new, 8 comparison, and 11 preredesign)
N of students = 682
N of pieces of student work = 1,378
*** $p < .001$, ** $p < .01$, * $p < .05$

Table B-2. Student Work Quality in Mathematics, Differences by School Type—HLM Estimates of Fixed Effects and Variance Components

	Model 1 (Baseline)			Model 2 (Full)		
Measure	**Coefficient**	**t-stat**	**Signif.**	**Coefficient**	**t-stat**	**Signif.**
Intercept	3.12	60.73	***	2.99	23.17	***
New (vs. comparison)	**-0.01**	**-0.14**		**0.01**	**0.02**	
Preredesign (vs. comparison)	0.38	6.3	***	0.35	5.05	***
Challenging	0.27	6.41	***	0.27	6.11	***
School risk index				0.01	0.32	
% lunch assistance				-0.14	-2.09	*
9th-grade test score				0.00	1.88	
Male				-0.09	-1.93	
Black				0.04	0.46	
Asian				0.10	1.08	
Hispanic/Latino				0.09	1.18	
Indian				-0.05	-0.19	
Variance components:						
Between student assignments	0.14	64%		0.15	65%	
Between students	0.08	36%		0.08	35%	
Total	0.21	100%		0.23	100%	

Note. N of schools = 31 (12 new, 8 comparison, and 11 preredesign)
N of students = 619
N of pieces of student work = 1,246
*** $p < .001$, ** $p < .01$, * $p < .05$

The Relationship Between the Quality of Student Work and Student 10th-Grade Achievement Scores

This analysis examines the relationship between the work that students do in class and the scores that they achieve on standardized student achievement tests. The results are reported in Tables B-3 and B-4. ELA data included 113 students who were taught by 16 teachers in 8 schools, while mathematics data included 92 students taught by 20 teachers in the same 8 schools.

The independent variable in this analysis is the quality of student work. To construct a student-level measure of student work quality and to do so while taking measurement errors of the scores into consideration, we used a three-level HLM latent variable model. The goal of the model below is to derive the random effects associated with students, i.e., u_{0k} , so they can be used as an independent variable of another analysis.

Level-1 Model (Measurement)

$$Y_{jk} = \pi_{jk} *(\text{DUMSWWT}) + \varepsilon_{jk}$$

Level-2 Model (Student Work)

$$\pi_{jk} = \beta_{0k} + r_{jk}$$

Level-3 Model (Students)

$$\beta_{0k} = \gamma_{00} + \gamma_{01}(\text{CHALLENGING}) + u_{0k}$$

This model specifies the MFRM score of a given piece of student work as consisting of the true score of the latent measure, represented by a dummy indicator variable (DUMSWWT), and a measurement error at level 1. Both the MFRM scores of student work and the dummy indicator were weighted by the inverse of the standard error of the measurement derived from the MFRM analysis, such that the residual variance at level 1 has a mean of zero and a variance of 1. In the student work model at level 2, the true score of the quality of a given piece of student work estimated at level 1 (π_{jk}) was modeled as a random effect, varying randomly across pieces of student work linked to a particular teacher assignment. The intercept of the level-2 model (β_{0k}), which represents the average level of student work quality for all the students, was further modeled at level 3 (student level) as a random effect and as a function of assignment type (challenging vs. typical). The random term u_{0k} corresponds to the average of student work quality, adjusted for precision in measurement, as well as the challenge variable. Using this as the independent variable, we next ran the following model to estimate the relationship between student work and student achievement.

Level-1 Model (Students)

$$Y_{ij} = \beta_{0j} + \beta_{1j}(\text{PRIORACH}) + \beta_{2j}(\text{STUDENT WORK}) + \beta_{3j}(\text{MINORITY}) + r_{ij}$$

where

- The outcome Y_{ij} is the achievement score.
- PRIORACH is a prior-year achievement score. It was centered around its grand mean.
- STUDENT WORK is a measure of student work quality.
- MINORITY is a dummy variable for student race being either Hispanic/Latino or Black. Other race/ethnicity categories are the reference group. All these student-level control variables were grand-mean centered.
- β_{0j} is the mean student achievement score for the classroom of teacher j.
- β_{1j} represents the effect of prior achievement score in the classroom of teacher j.
- β_{2j} represents the effects of student work in the classroom of teacher j.
- β_{3j} is the effect of students' ethnicity in the classroom of teacher j.
- r_{ij} is a student-level random error.

Level-2 Model (Teachers)

$\beta_{0j} = \gamma_{00} + \gamma_{01}$(Classroom-Mean PRIORACH) + γ_{02}(ZRISK) + u_{0j}

$\beta_{gj} = \gamma_{g0}$, g = 1~3.

- Classroom-mean PRIORACH is the class average prior-year achievement. It was centered around its grand mean.
- ZRISK is a school risk index based on the percentage of minority students and the percentage of students eligible for free or reduced-price lunch. It was centered around its grand mean.
- γ_{00} is the grand mean of the student achievement across all teachers/classrooms, adjusted for the student-level predictors.
- γ_{02} is the effect of the school risk index.
- u_{0j} represents the unique effect of teacher j on the classroom average student achievement.
- γ_{g0}, g = 1~3 indicates the fixed effects of the level-1 predictors across all teachers.

Table B-3. Relationship Between ELA Student Work Quality and 10th-Grade Reading Test Scores—HLM Estimates of Fixed Effects and Variance Components

	Model 1 (Baseline)			Model 2 (Full)		
Measure	**Coefficient**	**t-stat**	**Signif.**	**Coefficient**	**t-stat**	**Signif.**
Intercept	0.83	0.49		-0.01	-.06	***
Student work quality	**0.24**	**2.66**	**	**0.26**	**3.39**	**
9th-grade test score				0.48	5.33	***
Mean class 9th-grade score				0.30	2.69	**
Minority				0.14	0.97	
School risk index				0.04	0.37	
Variance components:						
Between teachers	0.44	46%		0.05	11%	
Within teacher	0.51	54%		0.40	89%	
Total	0.95	100%		0.45	100%	
N of teachers	16			16		
N of students	113			76		

Note. N of schools with 10th-grade test score data = 8
N of students = 113
*** $p < .001$, ** $p < .01$, * $p < .05$

Table B-4. Relationship Between Mathematics Student Work Quality and 10th Grade Mathematics Test Scores—HLM Estimates of Fixed Effects and Variance Components

	Model 1 (Baseline)			Model 2 (Full)		
Measure	**Coefficient**	**t-stat**	**Signif.**	**Coefficient**	**t-stat**	**Signif.**
Intercept	-0.01	-0.05		0.01	0.08	***
Student work quality	**0.14**	**1.39**		**0.08**	**0.85**	
9th-grade test score				0.41	4.31	***
Mean class 9th-grade score				0.29	2.08	
Minority				-0.05	-0.20	
School risk index				-0.05	-0.35	
Variance components:						
Between teachers	0.38	38%		0.08	13%	
Within teacher	0.61	62%		0.50	86%	
Total	0.99	100%		0.57	100%	
N of teachers	20			20		
N of students	92			92		

Note. N of schools with 10th-grade test score data = 8
N of students = 92
*** $p < .001$, ** $p < .01$, * $p < .05$

Differences in Students' Attitudes Toward Schooling Across School Types

Given the nested nature of our survey data (i.e., students nested within schools), we used HLM to examine the differences in student attitudes between different types of schools. The sample for the HLM analyses consisted of 45 schools in total. To construct the analytic sample, we first pooled survey data for 20 third-year new schools from the 2004 and 2005 data collections—i.e., we combined the 2004 data for 8 new schools that opened in 2002 and the 2005 data for 12 new schools that opened in 2003 for our analyses. Similarly, we pooled data for 19 preredesign schools surveyed between 2002 and 2005 and combined the data from these 19 schools with those from 11 comparison schools surveyed in 2004 and 2005 to form the group of 30 large high schools. Data for the 5 model schools were from the 2002 survey. Of the above 55 schools, 2 new schools, 4 preredesign schools, and 4 comparison schools were excluded from our analyses because of inadequate response rates and/or missing data on important measures. The final analytic sample for the cross-school-type comparisons included 18 third-year new schools, 22 large high schools, and 5 model schools.

We performed a separate analysis for each type of student attitude using a two-level HLM model. The level-1 model is a student-level model, which predicts a student's attitude with a set of student characteristics. The intercept of the level-1 model (β_{0j}), which represents the average level of student attitude for a given school, adjusted for student characteristics, is then used as the outcome in the school-level model at level 2, and modeled as a function of certain school characteristics. The specification of the HLM model is as follows:

Level-1 Model (Students)

$$Y_{ij} = \beta_{0j} + \beta_{1j} *(GRADE) + \beta_{2j} *(FEMALE) + \beta_{3j} *(NONENG) + \beta_{4j} *(MOMCOLLG) + \beta_{5j} *(MULTIOTH) + \beta_{6j} *(ASIAN) + \beta_{7j} *(HISPAN) + \beta_{8j} *(BLACK) + r_{ij}$$

Level-2 Model (Schools)

$$\beta_{0j} = \gamma_{00} + \gamma_{01} *(ZRISK) + \gamma_{02} *(LARGE) + \gamma_{03} *(MODEL) + \gamma_{04} *(YEAR03) + \gamma_{05} *(YEAR04) + \gamma_{06} *(YEAR05) + u_{0j}$$

$$\beta_{gj} = \gamma_{g0}, g = 1\sim8.$$

where

- GRADE, FEMALE, NONENG, and MOMCOLLG are grand-mean-centered variables representing the grade level, gender, English language learner status, and mom's education (whether mom had at least some college education) of student i in school j.
- MULTIOTH, ASIAN, HISPAN, and BLACK are a grand-mean-centered dummy variables for student race, with whites being the reference group.
- ZRISK is a grand-mean-centered variable measuring the extent to which school j enrolled disadvantaged student populations. It is based on the percentage of minority students and percentage of students eligible for free or reduced-price lunch.
- LARGE and MODEL are uncentered dummy variables for school type, with new schools being the reference school type.
- YEAR03, YEAR04, and YEAR05 are uncentered dummy variables for the year of survey administration, with year 2002 being the reference year.
- β_{0j} is the average level of student attitude within school j.
- $\beta_{1j}\sim\beta_{8j}$ represents the effects of student characteristics on student attitude within school j.
- γ_{00} is the average level of student attitudes across all schools, adjusted for the student characteristics included in the level-1 model.
- $\gamma_{01}\sim\gamma_{06}$ are the effects of school-level variables on the average level of student attitude in each school, adjusted for the student characteristics included in the level-1 model and other school-level predictors in the level-2 model.
- γ_{g0}, g = 1~8, represents the fixed effect of a given student characteristic on student attitude across all schools, adjusted for the other student characteristics included in the level-1 model.
- r_{ij} and u_{0j} are random errors at the student level and the school level respectively.

The level-2 coefficients for school type, γ_{02} and γ_{03}, are of particular interest; they represent the difference in student attitude between large schools and new schools and the difference between model schools and new schools, respectively.

The Relationship Between the Implementation of Effective-School Attributes and Student Attitudes Toward Schooling

To explore the relationship between implementation of effective-school attributes and student attitudes toward schooling, we used a path analysis. We added the implementation index as a school-level predictor to the HLM model for cross-school-type comparisons presented above. While the student-level model remains the same, the school-level model now becomes:

Level-2 Model (Schools)

$$\beta_{0j} = \gamma_{00} + \gamma_{01} *(\text{ZRISK}) + \gamma_{02} *(\text{LARGE}) + \gamma_{03} *(\text{MODEL}) + \gamma_{04} *(\text{YEAR03}) + \gamma_{05} *(\text{YEAR04}) + \gamma_{06} *(\text{YEAR05}) + \gamma_{07} *(\text{IMPLEMENTATION INDEX}) + u_{0j}$$

$$\beta_{gj} = \gamma_{g0}, \; g = 1\sim8$$

The estimate of the school-level coefficient γ_{07} represents the effect of the implementation index on school average level of student attitude, controlling for student characteristics, the school risk index, school type, and year of survey administration. To compute the path coefficient of the implementation index, we converted the unstandardized coefficient γ_{07} to a standardized coefficient based on the standard deviations of the implementation index and the outcome measure. The path coefficients for the school type effects (γ_{02} and γ_{03}) were similarly computed.

To obtain the path coefficients associated with the effects of school type on the implementation index, we further ran a school-level regression analysis with the implementation index as the outcome. The regression model is specified as follows:

$$Y = \beta0 + \beta1*(\text{ZRISK}) + \beta2*(\text{LARGE}) + \beta3*(\text{MODEL}) + \beta4*(\text{YEAR03}) + \beta5*(\text{YEAR04}) + \beta6*(\text{YEAR05}) + r$$

The standardized regression coefficients, β_2 and β_3, are the path coefficients for the school type effects on the implementation index, which represent the difference in the implementation index between large schools and new schools and the difference between model schools and new schools, respectively.

Change in Student Attitudes in Redesigned Schools

For redesigned schools, we compared student attitudes in their second year of redesign and those in their preredesign year, using paired-sample t-tests. The results of those comparisons are presented in Table B-5.

Table B-5. Change in Student Attitudinal Outcomes From the Preredesign Year to the Second Year of Redesign in Redesigned Schools

Outcome	Preredesign	Redesign	Difference	Significance
Student engagement—Interest	-0.968	-0.719	-0.249	*
Student engagement—Persistence	-1.136	-0.817	-0.319	*
Academic self-concept	-0.713	-0.774	0.061	
Educational aspirations	0.005	-0.089	0.093	

Note. N of redesigned schools = 27
N of preredesign schools = 8
* $p < .05$

Appendix C. Analysis for Chapter 3: Making Instruction Rigorous and Relevant

Rigor and Relevance of Teacher Assignments in New Schools and Comparison Schools

Although the focus of chapter 3 is on the comparison between new schools and comparison schools, we included data from other schools in our analyses to increase the precision of the parameter estimates. Specifically, the analytic sample for the cross-school-type comparisons of the rigor and relevance of teacher assignments included teachers in:

- Twelve new schools
- Eight comparison schools
- Eleven preredesign schools

The analytic sample for ELA consisted of 89 teachers and 647 assignments, and 81 teachers and 593 mathematics assignments.

Given the nested structure of the teacher assignment data (i.e., assignments are nested within teachers), we used hierarchical linear modeling (HLM) as the primary analytic method. Further, since the quality of teacher assignments is not directly observable and hence inevitably measured with error, we treated assignment rigor and relevance scores generated from the Many-Facet Rasch Measurement (MFRM) analyses as latent variables, taking advantage of the fact that MFRM provides both measures of assignment rigor and relevance and the standard errors of these measures.

Specifically, we constructed a three-level HLM latent variable model, where level 1 is the measurement model, level 2 is the assignment level, and level 3 is the teacher level. The purpose of the measurement model is to explicitly take into account the measurement errors in the MFRM scores of rigor and relevance, which were distinguished by two dummy indicator variables. The measurement model partitions the scores for each teacher assignment into a true score of the underlying latent measure (rigor or relevance) and a measurement error. The true values of the latent variables estimated at level 1 were then used as the outcomes at level 2 (assignment level) and modeled as a function of assignment type. The level-2 intercepts, which represent the average levels of rigor and relevance for each teacher adjusted for assignment type, were further modeled as random effects predicted by a set of teacher and school characteristics in the teacher-level model at level 3. Separate analyses were conducted for ELA and mathematics. The specification of the three-level model is as follows:

Level-1 Model (Measurement)

$$Y_{ijk} = \pi_{1jk} *(\text{DUMRELWT}) + \pi_{2jk} *(\text{DUMRIGWT}) + \varepsilon_{ijk}$$

where

- The outcome Y_{ijk} is the observed measure of latent variable i for assignment j given by teacher k, with the two latent variables being the levels of rigor and relevance of teacher assignment i (1 = rigor, 2 = relevance). The observed measures were weighted by the inverse of the standard errors of the measurement derived from the MFRM analysis.
- DUMRIGWT and DUMRELWT are two dummy variables indicating the specific latent variable that generated the observed MFRM score. The two dummies were un-centered at level 1 and weighted by the inverse of the standard errors of the observed outcome.
- π_{1jk} and π_{2jk} are the true scores of the two latent variables for assignment j given by teacher k.
- ε_{ijk} is a measurement error embedded in the observed score; $\varepsilon_{ijk} \sim N(0,1)$, given the weighting of both the dependent and independent variables.

Level-2 Model (Assignments)

$$\pi_{1jk} = \beta_{10k} + \beta_{11k}\,(\text{CHALLENGING}) + r_{1jk}$$
$$\pi_{2jk} = \beta_{20k} + \beta_{21k}\,(\text{CHALLENGING}) + r_{2jk}$$

where

- CHALLENGING is a dummy variable indicating the type of teacher assignment (1 = challenging, 0 = typical). It was grand mean centered and fixed at level 3.
- β_{10k} and β_{20k} are the mean rigor score and relevance score, respectively, for teacher k, adjusted for the type of assignment (i.e., challenging vs. typical).
- r_{1jk} and r_{2jk} are the random effects associated with assignment j given by teacher k on the rigor and relevance of assignment, respectively.

Level-3 Model (Teachers)

$$\beta_{10k} = \gamma_{100} + \gamma_{101}\,(\text{NEW}) + \gamma_{102}\,(\text{PREREDESIGN}) + \gamma_{103}\,(\text{ZRISK}) + \gamma_{104}\,(\text{NCE\_Z}) + u_{10k}$$

$$\beta_{11k} = \gamma_{110}$$

$$\beta_{20k} = \gamma_{200} + \gamma_{201}\,(\text{NEW}) + \gamma_{202}\,(\text{PREREDESIGN}) + \gamma_{203}\,(\text{ZRISK}) + \gamma_{204}\,(\text{NCE\_Z}) + u_{20k}$$

$$\beta_{21k} = \gamma_{210}$$

where

- NEW and PREREDESIGN are uncentered dummy variables representing new schools and preredesign schools, respectively, with comparison schools being the reference school type.
- NCE_Z is a grand-mean-centered measure of average ninth-grade achievement in ELA or mathematics of the class taught by teacher k. Given that different schools often use different standardized achievement tests, we converted ninth-grade test scores to NCEs, where possible, so that the class average prior achievement scores based on different tests were comparable. We also imputed

missing test scores via best-subsets regression using student-level eighth-grade test scores, ethnicity, sex, English language learner status, and free or reduced-price lunch status, where available.

- ZRISK a grand-mean-centered school risk index measuring the extent to which schools enrolled disadvantaged student populations. It was based on the percentage of students receiving free or reduced-price lunch and the percentage of underrepresented minority students.
- γ_{100} and γ_{200} are the grand means of the rigor and relevance of teacher assignment, respectively, across all teachers, adjusted for assignment type.
- γ_{110} and γ_{210} are the mean differences in rigor and relevance, respectively, between challenging teacher assignments and typical assignments across all teachers.
- γ_{101} and γ_{102} represent the differences in rigor of teacher assignments between new schools and comparison schools and between preredesign and comparison schools, respectively, adjusted for the independent variables in the model. Similarly, γ_{201} and γ_{202} represent the school type effects on relevance.
- u_{10k} and u_{20k} are the random effects associated with teacher k on the rigor and relevance of assignment, respectively, adjusted for the independent variables in the model.

By removing the school type variables (NEW and PREREDESIGN) from the above teacher-level model, we were able to obtain empirical Bayes estimates of assignment rigor and relevance for each teacher, adjusted for assignment type, class prior achievement, and the school risk index (not adjusted for school type). These teacher-level estimates were then aggregated to the school level and standardized to produce the rigor and relevance measures for individual schools.

Tables C-1 and C-2 report the results of the student work quality models. Model 1 is a baseline model, and Model 2 tests whether the relationships found for student work quality and student achievement hold up with a set of statistical controls.

Table C-1. ELA Rigor and Relevance: Differences by School Type—HLM Estimates of Fixed Effects and Variance Components

Measure	Model 1			Model 2		
	Coefficient	t-stat	Signif.	Coefficient	t-stat	Signif.
Rigor						
Intercept	5.16	23.84	***	5.23	23.78	***
New (vs. Comparison)	0.65	0.25		0.40	0.17	
Pre-redesign (vs. Comparison)	-0.20	-0.76		-0.36	-1.25	
Challenging	1.16	9.41	***	1.16	9.48	***
School Risk Index				-0.07	-0.62	
Mean class 9th-grade score				0.15	1.46	
Variance Components:						
Between Teachers	0.40	21%		0.38	20%	
Within Teacher	1.53	79%		1.52	80%	
Total	1.93	100%		1.90	100%	
Relevance						
Intercept	3.68	26.58	***	3.76	24.14	***
New (vs. Comparison)	0.71	3.23	**	0.70	3.29	**
Pre-redesign (vs. Comparison)	-0.42	-2.5	*	0.58	-3.04	**
Challenging	0.59	4.63	***	1.16	4.61	***
School Risk Index				-0.34	-3.76	***
Mean class 9th-grade score				-0.09	-0.89	
Variance Components:						
Between Teachers	0.39	68%		0.30	63%	
Within Teacher	0.18	32%		0.18	37%	
Total	0.57	100%		0.48	100%	

Note. N of schools = 31 (12 new, 8 comparison, and 11 preredesign)
N of teachers = 89
N of assignments = 647
*** $p < .001$, ** $p < .01$, * $p < .05$

Table C-2. Mathematics Rigor and Relevance: Differences by School Type—HLM Estimates of Fixed Effects and Variance Components

	Model 1			Model 2		
Measure	**Coefficient**	**t-stat**	**Signif.**	**Coefficient**	**t-stat**	**Signif.**
Rigor						
Intercept	3.68	18.66	***	3.68	18.7	***
New (vs. Comparison)	0.56	2.27	*	0.57	2.26	*
Pre-redesign (vs. Comparison)	0.55	2.24	*	0.53	2.11	*
Challenging	0.99	6.78	***	0.99	6.77	***
School Risk Index				-0.12	-0.93	
Mean class 9th-grade score				-0.08	-0.72	
Variance Components:						
Between Teachers	0.39	28%		0.37	27%	
Within Teacher	0.99	72%		0.99	73%	
Total	1.38	100%		1.36	100%	
Relevance						
Intercept	3.09	18.09	***	3.07	16.85	***
New (vs. Comparison)	0.90	2.74	**	0.92	2.77	**
Pre-redesign (vs. Comparison)	-0.21	-0.98		-0.17	-0.7	
Challenging	1.16	7	***	1.16	7	***
School Risk Index				-0.26	-1.53	
Mean class 9th-grade score				-0.25	-2.6	*
Variance Components:						
Between Teachers	0.84	43%		0.78	41%	
Within Teacher	1.11	57%		1.11	59%	
Total	1.95	100%		1.89	100%	

Note. N of schools = 31 (12 new, 8 comparison, and 11 preredesign)
N of teachers = 81
N of assignments = 593
*** $p < .001$, ** $p < .01$, * $p < .05$

The Relationship Between the Rigor and Relevance of Teacher Assignments and the Quality of Student Work

We used a two-step approach to assess the relationships between the rigor and relevance of teacher assignments and the quality of student work. We first constructed a three-level HLM latent variable model (Model 1) to derive estimates of assignment-level student work quality and then linked those estimates to measures of assignment rigor and relevance in a second three-level HLM latent variable model (Model 2), which produced the correlations among the three measures based on estimates of variance components at both the assignment level and the teacher level.

Specifically, Model 1 specifies the MFRM score of a given piece of student work as consisting of the true score of the latent measure, represented by a dummy indicator variable (DUMSWWT), and a measurement error at level 1. Both the MFRM scores of student work and the dummy indicator were weighted by the inverse of the standard error of the measurement derived from the MFRM analysis, such that the residual variance at level 1 has a mean of zero and a variance of 1. In the student work model at level 2, the true score of the quality of a given piece of student work estimated at level 1 (π_{jk}) was modeled as a random effect, varying randomly across pieces of student work linked to a particular teacher assignment. The intercept of the level-2 model (β_{0k}), which represents the average level of student work quality for all pieces of student work linked to a given assignment, was further modeled at level 3 (assignment level) as a random effect and as a function of assignment type (challenging vs. typical). Model 1 in an equation form is as follows. Model 1 is a three-level HLM latent variable model for deriving estimates of an assignment-level measure of student work quality:

Level-1 Model (Measurement):

$$Y_{jk} = \pi_{jk} * (\text{DUMSWWT}) + \varepsilon_{jk}$$

Level-2 Model (Student Work):

$$\pi_{jk} = \beta_{0k} + r_{jk}$$

Level-3 Model (Assignments):

$$\beta_{0k} = \gamma_{00} + \gamma_{01}(\text{CHALLENGING}) + u_{0k}$$

On the basis of the above model, we computed the level of student work quality for each assignment, adjusted for assignment type, as the sum of empirical Bayes estimates of the level-3 intercept and the level-3 residual (i.e., $\gamma_{00} + u_{0k}$). This assignment-level measure of student work quality was subsequently linked to the rigor and relevance of teacher assignments in Model 2. Model 2 is a three-level HLM latent variable model for assessing the relationships between the rigor and relevance of teacher assignments and student work quality.

In Model 2, the assignment-level measure of student work quality derived from Model 1 and the rigor and relevance of teacher assignments were treated as three latent outcomes, each consisting of a true score and a measurement error. To enable simultaneous estimation of the three latent outcomes, we stacked assignment data and student work data in such a way that each assignment occupies three rows in the data set, one for each latent outcome (rigor, relevance, and assignment-level student work quality). The specification of Model 2 at each level is as follows:

Level-1 Model (Measurement)

$$Y_{ijk} = \pi_{1jk} (DUMRELWT) + \pi_{2jk} (DUMRIGWT) + \pi_{3jk} (DUMSWWT) + \varepsilon_{ijk}$$

Level-2 Model (Assignments)

$$\pi_{1jk} = \beta_{10k} + \beta_{11k} (CHALLENGING) + r_{1jk}$$

$$\pi_{2jk} = \beta_{20k} + \beta_{21k} (CHALLENGING) + r_{2jk}$$

$$\pi_{3jk} = \beta_{30k} + r_{3jk}$$

Level-3 Model (Teachers)

$$\beta_{10k} = \gamma_{100} + \gamma_{101} (ZRISK) + \gamma_{102} (NCE\_Z) + u_{10k}$$

$$\beta_{11k} = \gamma_{110}$$

$$\beta_{20k} = \gamma_{200} + \gamma_{201} (ZRISK) + \gamma_{202} (NCE\_Z) + u_{20k}$$

$$\beta_{21k} = \gamma_{210}$$

$$\beta_{30k} = \gamma_{300} + \gamma_{301} (ZRISK) + \gamma_{302} (NCE\_Z) + u_{30k}$$

In the measurement model at level 1, the three latent outcomes are distinguished by three dummy indicator variables: DUMRIGWT, DUMRELWT, and DUMSWWT. Both the outcome and the dummy indicator variables were weighted by the inverse of the standard error of the measurement. At level 2, the true scores of the rigor and relevance of teacher assignments (π_{1jk} and π_{2jk}) and the quality of student work (π_{3jk}) were set to vary randomly across assignments within teachers. Assignment type, CHALLENGING, was used as a predictor for both the rigor and the relevance measures, but not for the student work measure because the assignment-level student work measure was already adjusted for assignment type in Model 1. The level-2 intercepts (β_{10k}, β_{20k}, and β_{30k}), which represent the average levels of rigor and relevance of teacher assignments and the quality of student work for each teacher, were further modeled as random effects and as a function of the school risk index (ZRISK) and classroom average prior achievement (NCE_Z) at level 3. The above model generated the correlations between the three latent outcomes at both the assignment level and the teacher level. The assignment-level correlations are presented in Table C-3.

Table C-3. Correlations Between Assignment Rigor, Relevance, and Student Work Quality at the Assignment Level in English/Language Arts and Mathematics

	Teacher Assignment Rigor	Teacher Assignment Relevance	Student Work Quality
English/language arts			
TA Rigor	1.00		
TA Relevance	0.43	1.00	
SW Quality	0.87	0.81	1.00
Mathematics			
TA Rigor	1.00		
TA Relevance	0.74	1.00	
SW Quality	0.78	0.56	1.00

Tables C-4 and C-5 present the HLM results for the relationships between assignment rigor and relevance and student work quality. Model 1 estimated the relationships with adjustment only for assignment type, whereas Model 2 incorporated additional adjustment for classroom average ninth-grade achievement and the school risk index.

Table C-4. The Relationship Between Assignment Rigor/Relevance and Student Work Quality in English/Language Arts—HLM Estimates of Fixed Effects and Variance Components

	Model 1			Model 2		
Measure	**Coefficient**	**t-stat**	**Signif.**	**Coefficient**	**t-stat**	**Signif.**
Intercept	0.62	0.91		0.81	1.33	
Rigor	**0.32**	**1.94**		**0.29**	**1.9**	
Relevance	**0.46**	**3.96**	***	**0.46**	**4.34**	***
School risk index				-0.18	-2.62	**
Mean class 9th-grade score				0.11	1.5	
Variance components:						
Between teachers	0.40	33%		0.31	28%	
Within teacher	0.81	67%		0.79	72%	
Total	1.21	100%		1.10	100%	

Note. N of schools = 31 (12 new, 8 comparison, and 11 preredesign)
N of teachers = 89
N of assignments with student work = 717
*** $p < .001$, ** $p < .01$, * $p < .05$

Table C-5. The Relationship Between Assignment Rigor/Relevance and Student Work Quality in Mathematics—HLM Estimates of Fixed Effects and Variance Components

Measure	Model 1			Model 2		
	Coefficient	t-stat	Signif.	Coefficient	t-stat	Signif.
Intercept	2.72	11.87	***	2.68	12.68	***
Rigor	**0.24**	**3.41**	**	**0.20**	**3.03**	**
Relevance	**-0.10**	**-2.45**	*	**-0.04**	**-1.03**	
School risk index				-0.00	-0.08	
Mean class 9th-grade score				0.10	3.27	**
Variance components:						
Between teachers	0.03	23%		0.01	12%	
Within teacher	0.10	77%		0.10	88%	
Total	0.13	100%		0.11	100%	

Note. N of schools = 31 (12 new, 8 comparison, and 11 preredesign)
N of teachers = 81
N of assignments with student work = 606
*** $p < .001$, ** $p < .01$, * $p < .05$

Appendix D. Analysis for Chapter 4: Sustaining School and System Change

Data Collection Methods

Data for the analyses presented in this chapter were collected from 17 districts:

- Anderson Union, CA
- Baltimore, MD
- Chicago, IL
- Cincinnati, OH
- Cotati-Rohnert Park, CA
- Denver, CO
- El Dorado, CA
- Milwaukee, WI
- New York City, NY
- Oakland, CA
- Oklahoma City, OK
- Providence, RI
- Ravenswood, CA
- Sacramento, CA
- St. Paul, MN
- West Clermont, OH
- Worcester, MA

Data were collected for individual schools and then aggregated at the district level. Data for each school were collected from national databases, state education Web sites, district Web sites, and school Web sites. The data elements reflect national, state, district, and school records as of the 2004-05 school year and include:

1. **High school name.** For each district, lists of high schools were obtained through the publicly available National Center for Education Statistics online database, the GreatSchools.net database, and state education Web sites to ensure that all high schools operating in the 2004-05 school year were included. Where the high school lists were inconsistent, the state education Web sites were considered authoritative because they had the most recent data. The national databases identified previously comprehensive schools that had been redesigned (i.e., broken into small learning communities, small autonomous schools, or academies within the same school building).

2. **Enrollment of students in grades 9-12.** Student enrollments from the 2004-05 school year were collected from state education databases. For schools where the grade levels span a wider grade range than 9-12, only 9th- through 12th-grade students were included in the enrollments. For Baltimore, Cincinnati, Providence, and St. Paul schools, enrollment figures were not available by individual grade levels,

so 9th- through 12th-grade enrollments were estimated by assuming an equal distribution of students across all grade levels.

3. **Type of school choice.** Each school in the 17 districts was determined to fall under one of five categories that describe the enrollment choices offered to students in 2004-05. The school categorizations were based primarily on the information presented on the district and school Web sites. Data collected during the case study visits were used when the choice data were not available online. The five categories of school choice are:

   - School where students can choose to attend and there are no admission criteria. These schools include charter, contract, magnet, and other choice schools.
   - Campus where students can choose to attend and additionally select the small learning community, small autonomous school, academy, house, or other small unit on the campus.
   - Campus that has a neighborhood attendance zone but in which students can select the small learning community, small autonomous school, academy, house, or other small unit on campus.
   - Comprehensive school with neighborhood attendance zone.
   - Other special type of school. The category includes alternative schools, schools with selective admission policies, and schools that do not fall under any of the other four categories.

4. **Enrollment of students in individual small learning communities.** Student enrollments from the 2004-05 school year were obtained from state, district, and school Web sites where available. If these data were not available online, students were assumed to be equally distributed across the small learning communities (SLCs) to estimate enrollment. In very few cases, comprehensive schools had a few SLCs or academies, but the majority of the students attended the larger comprehensive school. In these cases, student enrollment was included in the larger comprehensive school.

5. **Opening year for new schools of choice or year of redesign for schools with SLCs.** Opening and redesign dates for choice schools were obtained from school Web sites and conversations with school administrators.

6. **Type of curricular emphasis.** Each school's curricular emphasis was also determined from the information available on the school Web site. If no curricular emphasis was mentioned online, the school was assumed have a conventional instructional model. The categories of curricular emphasis are:

   - Theme-based instructional model. These schools organize instruction around curricular themes such as science and technology or social justice, community-related themes, or types of careers.

- Instructional model based on particular educational philosophies. These schools build their programs around particular instructional models, like student-centered instruction, or follow particular educational philosophies, such as open schooling. Some theme-based schools also have educational philosophies; these schools were not included in this curricular category.
- Conventional instructional model. These schools offer students a comprehensive curriculum with no specialized area of focus.

Data Analysis Methods

District and school data were summarized by using simple descriptive statistics.

Teacher Turnover Rates

Data Collection

Teacher rosters were collected from 12 new schools that were in the third or fourth year of operation in school year 2004-05. School rosters were compared across academic years and the numbers of faculty members who were retained or newly hired and those who left in 2002-03, 2003-04, and 2004-05 were cataloged. Data for third- and fourth-year schools were combined so that the percentage of teachers who stayed on staff, left, and came on board as schools moved from their first to second and second to third years could be calculated.

Data Analysis Methods

Teacher turnover data were summarized by using simple descriptive statistics.