EVALUATION OF
THE NEW CENTURY HIGH SCHOOLS INITIATIVE

Report on the Third Year

June 2006

Prepared for New Visions for Public Schools

By
Policy Studies Associates, Inc.
1718 Connecticut Avenue, NW
Suite 400
Washington, DC 20009

Executive Summary

Through New Visions for Public Schools and a partnership of grant makers, the New York City Department of Education has, over three years beginning in September 2002, opened 75 New Century High Schools (NCHS). This study by Policy Studies Associates describes approaches to instruction and patterns of achievement in those schools over that time period.

The evaluation's primary question is: Did NCHS schools promote academic success among students? We found strong indications that they did. In 2004-05, 77 percent of sampled NCHS eleventh-grade students were prepared for on-time graduation, which is a promising statistic in light of the 54 percent overall graduation rate in New York City public schools in 2003-04.

Student achievement in NCHS schools varied with changes in instructional conditions in the schools and especially with changes in the alignment of instruction with Regents requirements, the involvement of principals with academics, the appropriateness of student behavior, and the quality of support for classroom instruction. These conditions improved substantially across NCHS schools in 2003-04 and so did achievement. In 2004-05 these conditions settled a little lower and so did achievement.

Observations and interviews in NCHS schools confirmed, as survey data had suggested, that there were differences among NCHS schools with regard to the development and integration of instructional themes, the vitality of partnerships with external organizations, and the quality of student advisory periods. The fundamentals of good practice were, however, commonly recognized and embraced across NCHS schools (if not always realized).

Presented below is a digest of information gathered from student, teacher, principal, and community partner surveys, student records, and site visits to 26 NCHS schools.

Self-Reported Ratings of Effectiveness

Sponsors intended NCHS schools to be small (about 400 students each) and to evidence the following: rigorous instruction, personalized relationships, clear focus on teaching and learning, instructional leadership, school-based teacher-driven professional development, meaningful continuous assessment, community partners, family/caregiver partnership and involvement, youth participation and development, and effective uses of technology and resources.

Using surveys, we asked teachers, students, principals, and community partners about their schools' achievement of those qualities. The findings are

presented below organized around two time-ordered questions: Did first-year ratings improve for successive groups of NCHS schools? Did ratings improve within groups of NCHS schools over time? We expected that ratings would improve over time across school groups and within school groups. This is not, however, exactly what happened.

Did first-year ratings improve for successive groups of NCHS schools?

This question asked if successive groups of NCHS schools got off to a better start, with groups defined by the years in which the schools opened. In important ways, they did. Teachers provided higher initial ratings of the rigor of instruction and the classroom involvement of principals in schools that started *after* 2002-03, the year in which the first group of NCHS schools opened their doors. Student ratings tended, however, to be constant over time, including their ratings of instructional relevance and challenge and of relationships with peers and teachers. Students rated the quality of teacher assessments of student learning as quite high, but consistently so, over time. Student ratings of their opportunities to engage in leadership activities dipped after 2002-03.

Did ratings improve within groups of NCHS schools over time?

We asked if staff and student ratings of the educational environments in their schools changed over time within groups of schools. In general, unusually high and low ratings regressed to the mean (or gravitated to average) with time. Ratings of student discipline and principals' levels of classroom involvement, which were low in Group 1 schools in their first year, rose, and the exceptionally high first-year ratings of student discipline and principals' levels of classroom involvement in Group 2 schools fell. As one might expect given the addition of successive grade levels in NCHS schools each year and the resulting growth in the numbers of students and teachers, ratings of individual teacher influence, parent-outreach effort, and parent involvement sloped mildly downward over time.

Measurement of Student Outcomes

We framed analyses of student outcomes around three questions: What were the characteristics of students who were enrolled in NCHS schools in 2004-05? Were these students headed for on-time graduation? Did the achievement of NCHS students exceed the achievement of comparison-group students?

What were the characteristics of students enrolled in NCHS schools in 2004-05?

Relative to public high school students citywide, students in NCHS schools were more likely to be female, African American or Hispanic, and poor. They were less likely to be English Language Learners, new immigrants, or special education students. Lower proportions of NCHS students were proficient or advanced in English Language Arts or math upon entry into high school than were proficient or advanced citywide.

Were these students headed for on-time graduation?

Using credit accumulation as the measure for likelihood of high school graduation, 80 percent of students who originally enrolled in NCHS schools as ninth-graders were categorized as "on track" for graduation after one year, 68 percent were on-track after two years, and 77 percent were on-track after three years. Although it was a policy in some NCHS schools to postpone offering Regents exams, on average, as of 2004-05, NCHS students were passing Regents exams at a rate adequate for on-time graduation. By the end of tenth grade, the average NCHS student had passed 2.08 of the five Regents exams required for graduation. By the end of eleventh grade, the average NCHS student had passed 3.49 Regents exams.

Did the achievement of NCHS students exceed the achievement of comparison-group students?

Rates of school attendance, credit accumulation, and grade promotion were higher among NCHS students than among comparison-group students who were matched to NCHS students by means of propensity scoring. The differences were statistically significant. NCHS students also appeared to have higher school-persistence rates than did comparison-group students, although the differences were not statistically significant. With regard to Regents exams, we found that NCHS students in Group 1 schools (those initiated in 2002-03) were less successful than comparison-group students. NCHS students in Group 2 schools (those initiated in 2003-04) were, however, more successful than their matches. Suspension rates were higher among NCHS students in 2004-05 than among comparison-group students.

Assessments of Schools by Outside Observers

In site visits conducted in 2004-05 in 26 of 75 schools we explored three questions: Was the climate in NCHS schools safe and focused on instruction and youth development? Did advisory periods, curricular themes, partnerships, and

professional support—matters of particular interest to NCHS sponsors—contribute to effectiveness? What obstacles did NCHS schools face?

Were NCHS schools small, safe, and focused on instruction and youth development?

The climate in sampled NCHS schools was generally aligned with preferred practice in these domains. Almost without exception, the schools were safe, academically focused, and socially supportive. Seventy-five percent of the schools were rich with instructional supports and youth development activities.

Did small size, advisory periods, curricular themes, partnerships, and professional support—matters of particular interest to NCHS sponsors—contribute to effectiveness?

We found considerable diversity among the schools with regard to their thematic integration, teachers' assessment of professional development activities, and the quality of partnerships with nonprofit external organizations. Advisory periods, while sometimes helpful, were frequently under-developed both for lack of a unifying curriculum and for lack of staff training. The small size of the schools and the availability of supportive academic services after-school, on the weekend, and over the summer seemed, however, to contribute consistently to student engagement and learning.

What obstacles did NCHS schools face?

NCHS schools were not perfect. Suspension rates in NCHS schools grew from 2 percent in 2002-03 to 4 percent in 2003-04 to about 6 percent in 2004-05, a rate that equaled the citywide average. The optimistic interpretation of this trend is that higher suspension rates among NCHS students signaled particular vigilance on the part of NCHS staff members. An alternative view is that it signaled that the schools were experiencing growing problems. In either case, it's hard to see rising suspension rates as benign for the students involved. Generally speaking when suspension rates go up, the overall level of students' connectedness to school goes down (McNeely, Nonnemaker, & Blum, 2002).

Respondents identified several conditions that they believed created tensions in NCHS schools. Teachers and students both reported that social bonds were adversely affected with the addition of each successive grade level. It was not their view that all NCHS schools should have only 100 or 200 students, but it was their view that growth in numbers came with a price tag and that some high schools in the system should remain *very* small.

Like principals and teachers throughout the city school system, many NCHS staff members saw their schools as crowded. The perception of crowding was frequently associated with the presence of several schools in the same building. In many of these sites, there tended to be competition for resources among the schools, and staff and students reported regular fights.

The most sensitive contextual challenge for NCHS schools was a citywide practice known as over-the-counter admissions. This is a practice, according to school system officials, of assigning students who did not participate in the high school admissions process (or who requested a transfer due to documented medical or safety reasons or due to travel hardship) to schools with available seats based on geographic proximity and student interest. According to respondents, despite the school system's efforts to help students select appropriate available placements, a high proportion of students admitted over the counter were not, in fact, interested in NCHS themes or in the special routines sometimes followed in these schools, such as wearing school uniforms.

Conclusions

As useful as comparison-group studies are for estimating possible program effects, they have limits. Selection effects and other unmeasured factors can be at play in any comparison-group study. In this study, we found that NCHS students out-performed comparison-group students. We also found that more positive learning outcomes were related to implementation of proven instructional practices. It is this convergence of implementation and outcome findings (not the mere fact that NCHS students out-performed their matches) that persuades us that NCHS schools contributed importantly to the educational achievement of enrolled students.

Acknowledgements

New Visions for Public Schools sponsored this evaluation of the New Century High Schools Initiative. We are grateful to New Visions and to the schools in its network for the opportunity to evaluate their important work.

We are also grateful to the New York City Department of Education (DOE). DOE reviewed and approved the design of this research and provided the student-level data used in outcome analyses.

The findings and conclusions presented in this interim report of the evaluation were, of course, determined independently. They do not represent the official position of New Visions for Public Schools or the New York City Department of Education.

This report was the work of many individuals at Policy Studies Associates, Inc. Eileen Foley designed and oversaw analyses and wrote the report. Elizabeth Reisner directed the project and edited the report. Sara Allender managed the collection and analysis of survey data and the preparation of survey graphics. Marjorie Cohen trained interviewers and observers and managed the collection, analysis, and presentation of qualitative data. Allan Klinge analyzed student outcome data and prepared related exhibits. Dwayne Smith and Mark Wilson collected and analyzed survey and qualitative data. Kim Thomas prepared the report for publication.

Contents

LIST OF EXHIBITS

I. INTRODUCTION

Aiming to reinvigorate secondary education in New York City, a partnership of educators and grant makers opened 75 small high schools in school years September 2002-03 through September 2004-05. Known as the New Century High Schools Initiative, this partnership, managed by New Visions for Public Schools, included the New York City Department of Education (DOE), the United Federation of Teachers, the Council of Supervisors and Administrators, the Bill & Melinda Gates Foundation, the Carnegie Corporation of New York, and the Open Society Institute. Sponsors designed these new schools to be small (about 400 students each) and to reflect the following emphases:

- *Rigorous instructional program* enabling every student to master challenging skills, content, and state standards through relevant, individualized, in-depth, and inquiry-based teaching
- *Personalized relationships* between students and teachers, characterized by close continuous communication and each student having at least one adult to coordinate the support needed for the student to achieve postsecondary goals
- *Clear focus* on teaching and learning and expectations that every student will succeed
- *Instructional leadership* through effective collaboration and school-wide support for teaching and learning
- *School-based teacher-driven professional development and collaboration* that would be results-driven, standards-based, and embedded in the daily work of the school
- *Meaningful continuous assessment* to diagnose student needs and improve instruction
- *Community partners* that offer support and opportunities for students, families, and the school community and contribute significantly to the school's planning processes and operations
- *Family/caregiver partnership and involvement* in governance and the design of the school's education program
- *Youth participation and development*, characterized by student voice in teaching and learning and shared responsibility for the operation and governance of the school

- *Effective uses of technology and resources,* including print, visual, audio, and electronic technology

Sponsors of New Century High Schools (NCHS) recognized the importance of regular feedback to inform their work, the DOE's wider small schools initiative, and the broader public. New Visions asked Policy Studies Associates, Inc. (PSA) to conduct a comprehensive evaluation beginning in 2002-03. In tandem with NCHS stakeholders, we designed an evaluation from the sponsors' program theory; i.e. their assumptions about how NCHS would create social benefits (Rossi, Freeman, & Lipsey, 1999).

Exhibit 1.1 presents that program theory. Briefly stated, architects of the initiative saw effective implementation and positive results as contingent upon careful planning, direct support (especially in the form of professional development), the alignment of policy systems, and the collective deployment of community and school system resources.

During the initiative's first two years of operation, we examined the dimensions of implementation outlined in the program theory. Data sources included surveys completed annually by school-level stakeholders (principals, partners, teachers, and students) and on-site interviews in all of the 12 NCHS schools operating in 2002-03 and in 20 of the 30 schools operating in 2003-04.

Generally speaking, stakeholders viewed the schools' academic programs favorably. Students typically described their teachers as holding clear, high, and consistent expectations for learning. Teachers typically said the schools were focused and rigorous and that principals were effective and collaborative leaders.

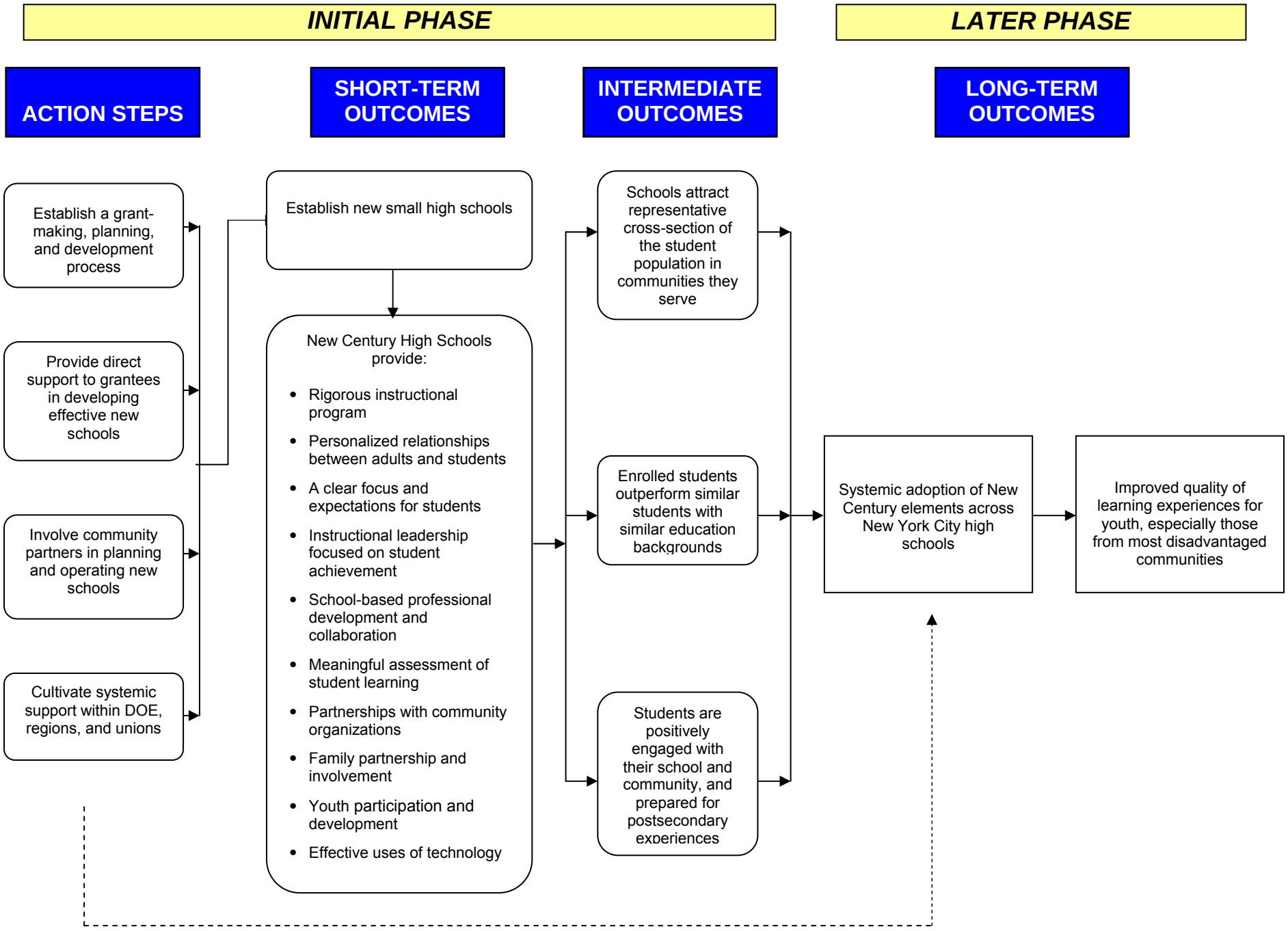
The schools were not, of course, picture-perfect. Parent participation was lower than expected. Some students acted out, and relationships among students on campuses shared by several schools were occasionally volatile.

Research Questions Guiding Year 3 Evaluation

This report presents findings based on the third year of NCHS implementation, school year 2004-05. Research activities in 2004-05 were primarily aimed at assessing educational outcomes of NCHS students. We used data on school operations and program implementation primarily to decide whether outcomes could reasonably be attributed to the initiative.

We framed three sets of questions to guide the 2004-05 data collection. The first set of questions asked if NCHS schools evidenced increasing or decreasing levels of the instructional and organizational qualities that sponsors valued. This was a two-part question. The first part inquired about start-up-year ratings *across* successive groups of NCHS schools (with groups determined by

Exhibit 1.1: Program Theory



the years schools opened). The second part inquired about ratings over time *within* groups of NCHS schools.

The second set of questions asked if NCHS students were more successful than comparison-group students on academic measures, including rates of persistence, attendance, credit accumulation, and Regents exams passed. We aligned ninth-grade outcome analyses with implementation analyses to look for patterns of association, and we studied the cumulative achievement of students who remained in the schools for two and three years.

The third set of question asked if the instructional environment in NCHS schools was aligned with NCHS program emphases and empirically validated best practices. We also asked about challenges facing the schools.

Limits of Outcome Assessment

To understand the performance of students in NCHS schools, we employed a quasi-experimental research design using a constructed comparison group as the counterfactual condition. A comparison group design was optimum for this study, given the infeasibility of assigning students at random to schools that were not over-subscribed.

While appropriate given the research circumstances, quasi-experimental designs nonetheless set limits on the certainty that can be attached to research findings. In the absence of random assignment, unmeasured characteristics in the treatment and comparison groups can affect outcomes. We have tried through statistical means to limit this possibility, but it continues to exist.

As a separate matter, we note that performance data had a clear hierarchical structure; treatment- and comparison-group students were nested within schools. The analytic strategy generally preferred under these circumstances is hierarchical linear modeling. But this was not the approach we adopted. Given the preliminary nature of this investigation, our priority was to explore the data carefully, and we did so using traditional statistics. Boundaries between schools were collapsed, and academic performance was explored at the student-level. We run some risk, therefore, of over-stating the statistical significance of differences between NCHS and comparison-group students.

Summary of Year 3 Findings

The central finding is that student achievement was enhanced by attendance in NCHS schools. Students in the NCHS research sample on average accumulated credits at a rate consistent with on-time graduation, and they out-performed comparison-group students.

Start-up year outcomes were best for the NCHS schools that opened in the second year of the initiative (2003-04), as compared to schools opening in the initiative's first and third years of operation. Teacher survey data aligned with outcome data. Teachers in Group 1 schools (those that opened in 2002-03) reported a lesser emphasis on academics in their start-up year than did teachers in later groups and years. According to survey data, Group 1 teachers re-oriented their approach in 2003-04, and according to outcome data, student performance improved in that year.

Information gathered in interviews with NCHS stakeholders indicated a less favorable context for instruction in 2004-05 than in 2003-04, which was a banner year for NCHS schools. According to principals, but not independently verified by us, in 2004-05, the quality and amount of space available to NCHS schools decreased relative to the schools' increasing needs and enrollment. Principals, teachers and students also said that fights were common between students in schools sharing the same campus. According to teachers staffing ratios were less favorable in 2004-05 than earlier.

In 2004-05, principals and teachers also reported a problem with "over-the-counter admissions." According to school system officials, this is a practice of assigning students who did not participate in the high school admissions process (or who requested a transfer due to documented medical or safety reasons or due to travel hardship) to schools with available seats based on geographic proximity and student interest. According to interview respondents, a high proportion of students admitted over the counter were not, in fact, interested in NCHS themes or in the special routines sometimes followed in these schools, such as wearing school uniforms.

Organization of Report

This report is presented in six chapters and one appendix. Chapter II is an overview of our research methods. Chapter III presents indexes we used to track program implementation across time. Chapter IV contrasts academic outcomes for NCHS and comparison-group students. Chapter V summarizes commentary from stakeholders about life in NCHS schools in 2004-05. Chapter VI discusses our findings and presents recommendations for program development and further research. The appendix elaborates methods used to distill research samples for student achievement analyses and discusses matters relating to the external validity of the study.

II. RESEARCH METHODS

This chapter briefly describes quantitative methods including sample selection processes, the uses we made of student records, and our statistical procedures. Qualitative methods are presented in Chapter V.

Sample Selection

Below we describe first the sample selection processes used in outcome analyses and next the sample selection procedures used in implementation analyses.

Outcomes

For achievement analyses, as earlier noted, we employed a quasi-experimental design with a constructed comparison group as the counterfactual. Comparison-group students were selected by means of propensity-score matching. First, we matched the 30 NCHS schools operating in 2003-04 with 12 larger traditional public high schools (capacity $\geq$ 750 students).[1] Next, we matched ninth-grade entrants to NCHS schools with ninth-grade entrants to comparison-group schools.[2]

We chose to concentrate on ninth-grade entrants to NCHS schools, based on the size and importance of this student group. From among the 12,634 entrants to NCHS schools in school-year 2002-03 through school-year 2004-05, 10,266 students enrolled as ninth-graders, meaning that more than 80 percent of enrolled students had entered NCHS schools as ninth-graders.

The academic records of a substantial number of these students (N = 2,731) lacked some data element required for propensity matching, leaving 7,535

[1] Schools were matched based on the following: the percent of their students eligible for free and reduced price lunch; the percent who were recent immigrants; the percent who were English Language Learners; the percent African American, Hispanic, Asian, White, and Other; and students' eighth-grade English Language Arts and math scale scores. None of the comparison group schools we selected was scheduled for phase-out.

[2] Students were matched based on eighth-grade academic records (attendance, math scale score, and English Language Arts scale score) and background characteristics (gender, race, age, recent immigrant status, special education eligibility, ELL eligibility, free or reduced-price lunch eligibility, and grade level). Age differences could not exceed 180 days.

eligible participants. An appropriate match[3] was found for 1,875 of these students,[4] and all were included in holding power analyses. From within this group, smaller samples were available for specific analyses. For example, 1,630 students were included in attendance analyses as they had the requisite attendance data.[5]

A strength of this research is the equivalence of NCHS and comparison-group students on observed variables. (See the appendix for information regarding the external validity of the research.) Exhibit 2.1 presents differences between the 1,875 NCHS and comparison-group students included in analyses of holding power using *t*-test procedures. Matches were exact for categorical variables like gender, race, free lunch status, etc. so chi-square test statistics are thus omitted. There were no statistically significant differences between the groups on measured characteristics.

Exhibit 2.1
Differences between NCHS and Comparison-Group Students in the Eighth Grade

Characteristic	NCHS Group (n = 1,875)		Comparison Group (n = 1,875)		Significance
	M	*SD*	*M*	*SD*	*t*
Attendance	0.931	0.045	0.931	0.044	1.068
Math Scale Score	705.540	21.125	705.450	21.134	0.836
English Scale Score	686.592	15.820	686.552	15.847	1.742
Age (in days)	4979.162	148.902	4976.925	147.500	1.107

*$p \leq .05$, **$p \leq .000$.

Exhibit reads: The eighth-grade school attendance rate for NCHS and comparison group students was 93.1%.

[3] Matches were considered appropriate under the following circumstances: (a) students were identical with regard to race, gender, recent immigrant status, ELL eligibility, free- or reduced-price lunch eligibility, and special education status, (b) attendance differences did not exceed 0.15 standard deviations, ELA scale score differences did not exceed 0.10 standard deviations, and math scale score differences did not exceed 0.20 standard deviations, and (c) age differences did not exceed 180 days.

[4] The odds of finding an appropriate match grew less favorable over time, for two reasons. First, comparison schools increased their rate of holding ninth-grade students over, from 39.0 percent in 2001-02 to 42.2 percent in 2002-03 to 45.3 percent in 2003-04. As holdovers grew in numbers, new admissions declined. Second, NCHS schools increased in number and enrollment. The ratio of NCHS to comparison group students was about 1:10 in 2002-03. In 2003-04, it was about 1:3. By 2004-05 there were only slightly more comparison group students than NCHS students, and the ratio was about 1:1.

[5] The appendix clarifies the extent to which propensity-score matching yielded research groups with comparable characteristics to the ninth-grade population with data.

Implementation

We successfully administered four surveys assessing program implementation in each of three years: a principal survey, a partner survey, a teacher survey, and a student survey. Response rates were high across administrations of these instruments. Annually over 85 percent of principals and partners responded, over 75 percent of teachers responded, and no less than 73 percent of students responded. Response rates were higher for schools that opened in 2002-03 and 2003-04 than for schools opening in 2004-05. In the first two years of the initiative, only one principal failed to return a principal survey; whereas seven of 45 principals joining the initiative in 2004 failed to return a principal survey.

Student Records

Student demographic and achievement data were obtained uniformly from the New York City Department of Education (DOE) in SPSS formatted files. For 70 percent of students, records were sufficiently complete to allow for propensity-score matching. It was particularly challenging to match students based on their eligibility for free- and reduced-price lunch.

Each year, parents of students in New York City public schools are asked to complete a school meal application (Form1041). This form is required for determining federal reimbursement for meals as well as for determining Title I funding for schools. A high proportion of parents fail to complete a 1041, perhaps because many know they are not eligible. As a result, data were often missing for this important variable. To address this problem, we matched eligible students with eligible students, ineligible students with ineligible students, and students for whom data were missing with students for whom data were missing.

Statistical Procedures

Quantitative data in this report are presented in tables (outcome data) and graphics (implementation data). When we label a difference as statistically significant (or unlikely to have occurred by chance), we refer consistently to the $p \leq 0.05$ level of significance.

In implementation analyses, to ease interpretation, we report confidence intervals for indexes in pictures (or figures). A vertical line extending above and below each mean represents the estimated 95 percent confidence interval for that mean. When confidence bands for two means do not overlap, the means are statistically different.

In outcome analyses, to clarify the practical significance of differences between NCHS and comparison-group students, we include effect sizes (in parentheses) as well as significance levels. Effect sizes were calculated in three steps. First, we derived the mean of differences in the paired performance of NCHS and comparison students. Next, we derived the standard deviations of those difference scores. And finally, we divided the mean difference by the standard deviation of the difference (d = M/SD). Many in the social sciences would agree that effect sizes of 0.2 or higher reference non-trivial differences between groups (Cohen, 1977). Effect sizes reported in this study frequently exceed that standard.

III. FINDINGS REGARDING IMPLEMENTATION

Surveys administered in all NCHS schools in 2002-03 through 2004-05 asked principals, partners, teachers, and students to gauge the degree to which their schools operated in conformity with NCHS founding principles. In analyses, we bundled respondents' answers to related survey questions into composite scores (or indexes) and analyzed change over time in these composite scores. We examined the patterns of change within schools that opened in the same year and between groups of schools that opened in different years. We label school groups as follows: Group 1 schools are those that opened in 2002-03, Group 2 schools are those that opened in 2003-04, and Group 3 schools are those that opened in 2004-05.

We developed cumulative scores, or indexes, in four steps. We first selected items based on their alignment with the construct under consideration (face validity). Next, we selected from each group those items with good response variation. We examined the resulting subset of items for internal consistency (Cronbach's alpha), and we then trimmed the list to items that optimized cohesion.

All items included in each implementation index carried the same weight—one point. When respondents affirmed that a condition aligned with an NCHS founding principle existed in their school, one point was awarded. Points were added across items and raters, and sums were divided by the number of items multiplied by the number of raters. This process placed all indexes, regardless of the numbers of items or the number of raters, on a scale that ranged from 0.00 to 1.00, with higher scores representing more favorable ratings. A score of .80 means that 80 percent of responses were affirmative; a score of .40 means that 40 percent of responses were affirmative.

Summary

We focused data analyses on two research questions: Did first-year ratings improve for successive groups of NCHS schools? Did ratings over time improve within groups of NCHS schools? A brief summary, organized by research question, is presented immediately below, and a detailed presentation follows thereafter.

Did First-Year Ratings Improve for Successive Groups of NCHS Schools?

This question asks if successive groups of schools got off to a better start. In important ways, they did. Teachers provided higher initial ratings of the rigor of instruction and the classroom involvement of principals in schools that opened *after* 2002-03. Student ratings tended to be constant over time including ratings of instructional relevance and challenge and ratings of relationships with peers and with teachers. Students rated the quality of teacher assessments of student learning as quite high, but consistently so, over time.

Did Ratings Improve within Groups of NCHS Schools over Time?

This question asks if ratings of schools changed over time. In general, ratings seemed to settle down over time. Students in Group 1 schools rated their first-year level of engagement in leadership activities substantially higher than did Group 2 and Group 3 students. But Group 1 student ratings dropped in 2003-04 and remained at that level in 2004-05. Teachers in Group 2 schools rated the instructional climate and classroom involvement of principals substantially higher than did Group 1 and Group 3 teachers. But Group 2 teacher ratings dropped in 2004-05. As one might anticipate, school size appeared to have substantial bearing on the power of schools as communities. With the addition of successive grade levels each school year, ratings of teacher influence, parent outreach, and parent involvement sloped mildly downward.

Findings by Index

In the sections that follow, findings are presented with added detail. Introductory tables identify the following: (1) select survey items that formed the implementation indexes, (2) respondents who provided data, and (3) index reliabilities in the form of Cronbach's alpha. Like other reliability coefficients, Cronbach's alpha ranges from 0.00 to 1.00. Scores toward the high end of the range (above 0.70) suggest that items in the index are measuring the same thing.

Following the description of each implementation index, two figures are presented. The first figure depicts *start-up year* ratings for successive groups of NCHS schools. The second figure depicts *patterns of performance over time* within groups of NCHS schools. Within these figures, circles represent index means and vertical lines above and below the circles represent the estimated 95 percent confidence intervals of those means. Confidence bands for some indexes were exceptionally small, and they are not visible in our graphics (see Exhibits 3.8, 3.10, 3.12, 3.24, 3.30, 3.32, 3.34, and 3.36). When confidence bands for two means do not overlap, the means are statistically different.

Rigorous Instructional Program

We distilled four indexes from survey data to gauge whether schools were instructionally rigorous. Each is presented below. The first index measures teachers' perceptions of the instructional usefulness of New York State Regents examinations.

Rigorous Instructional Programs enable every student to master challenging skills, content, and state standards through relevant, individualized, in-depth, and inquiry-based teaching.

Scale	Survey	Questions	Alpha
A. Content aligned with Regents Standards	Teacher	20a-e	.90

Were Regents standards useful for the following…

Q20a Selecting curricular materials
Q20b Developing curriculum
Q20c Designing classroom assessments
Q20d Developing a school improvement plan
Q20e Designing or selecting professional development opportunities

Exhibit 3.1 shows teachers' first-year ratings of the instructional usefulness of Regents requirements. Group 2 and Group 3 schools paid more attention to Regents in their start-up year than Group 1 schools. Exhibit 3.2 shows that changes over time within groups of schools were not statistically significant.

Exhibit 3.1
First-Year Use of Regents

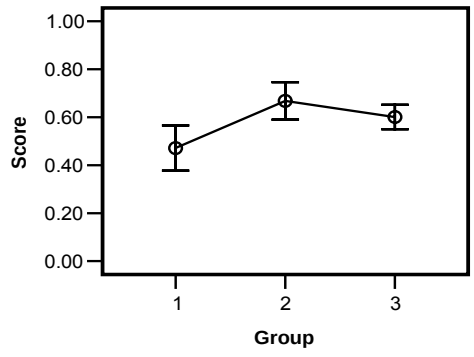
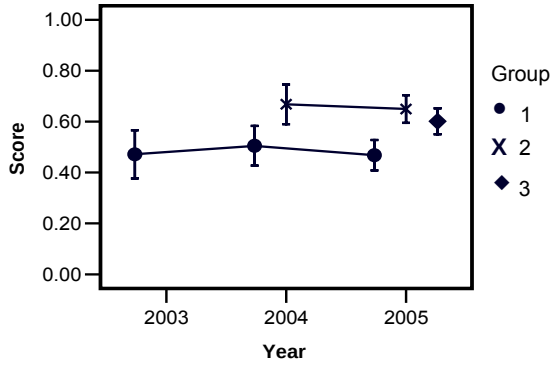


Exhibit 3.2
Use of Regents Over Time



The second index relating to instructional rigor was formed from a set of items concerned with student discipline. The scale is presented on the page to follow.

Rigorous Instructional Programs (continued)			
Scale	**Survey**	**Questions**	**Alpha**
C. Student discipline	Teacher	10 b, d,e,l,m,n,p,r, y	**.82**

Were the following problems in your school…

Q10b Student absenteeism
Q10d Students cutting class
Q10e Physical conflicts among students
Q10l Verbal abuse of teachers
Q10m Student disrespect of teachers
Q10n Student disrespect for other students
Q10p Student apathy
Q10r Lack of parental involvement

Exhibit 3.3 shows teachers' first-year ratings of the level of discipline in their schools. Group 2 teachers gave their schools statistically higher ratings than did teachers in other groups. Exhibit 3.4 shows that teachers' perceptions of the quality of discipline in Group 2 schools fell to the norm by 2004-05.

Exhibit 3.3
First-Year Discipline Level

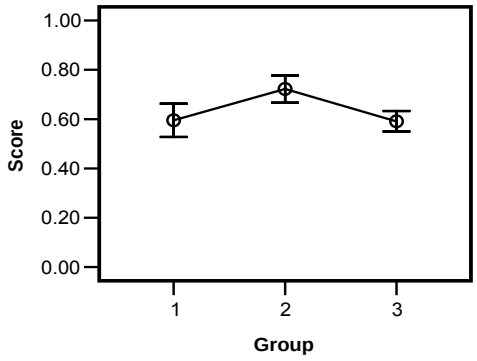
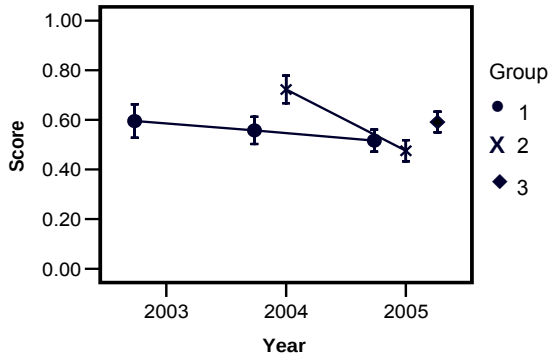


Exhibit 3.4
Discipline Level Over Time



The third index relating to instructional rigor focused on teachers' perceptions of the quality of support for instruction in their schools. The index is described on the page to follow and a sample of items are presented.

Rigorous Instructional Programs (continued)

Scale	Survey	Questions	Alpha
D. Support for instruction	Teacher	35a, c, d-k	.75

Do you agree with the following…

Q35a	Teachers in this school are evaluated fairly
Q35c	Teachers participate in making most of the important educational decisions in this school
Q35d	Teachers receive a great deal of support from parents for the work they do
Q35e	Necessary instructional materials are available as needed by the staff
Q35f	In this school, staff members are recognized for a job well done
Q35g	Parents, community partners, and community members share in school decision-making and governance
Q35h	Parents, community partners, and community members are involved in the educational program
Q35i	Teachers are continual learners and team members through professional development, common planning, and collaboration
Q35j	Technology is used in this school to manage curriculum, instruction, and student progress
Q35k	There are formal arrangements within the school that provide opportunities for teachers to discuss and critique their own and others' instruction

Exhibit 3.5 shows teachers' first-year ratings of the quality of support for instruction in their schools. Group 2 and Group 3 teachers gave their schools statistically higher ratings than Group 1 teachers. Exhibit 3.6 shows, however, that teachers' perceptions of the quality of support for instruction were virtually the same across school groups by 2004-05.

Exhibit 3.5
First-Year Instructional Support

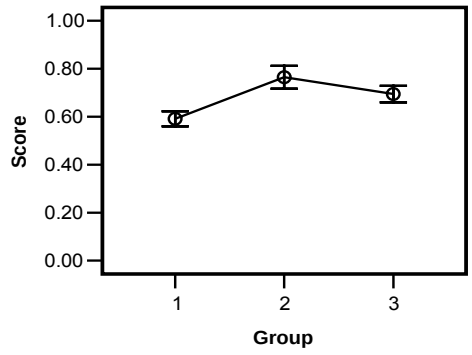
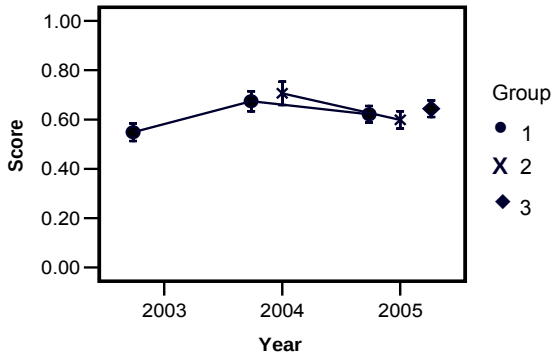


Exhibit 3.6
Instructional Support Over Time



Unlike the preceding indexes, which focused on teachers' perspectives, the final index in the instructional rigor series examined students' perceptions of instruction. The index is described on the page to follows.

Rigorous Instructional Programs (continued)			
Scale	**Survey**	**Questions**	**Alpha**
E. Classes interesting, challenging, relevant, etc.	Student	6a,b,d,f,g	.74

Would the following statements apply in most of your classes…

Q6a I learn a lot
Q6b I spend most of the time learning new things
Q6d I am challenged to work hard
Q6f What I am learning is interesting
Q6g The class teaches me how to apply things I've learned before to new situations

Exhibit 3.7 shows students' first-year ratings of instruction. Group 2 students gave their schools lower ratings than students in other groups. This contrasts with the statistically *higher* ratings that Group 2 teachers gave the same schools on matters relating to instructional climate and Regents preparation. This difference introduces the possibility that there is a tension between teachers' investments in Regents' preparation and students' experience of instruction as interesting. Exhibit 3.8 shows that student ratings of instruction remained consistent across school groups over time.

Exhibit 3.7
First-Year Students' Instructional Ratings

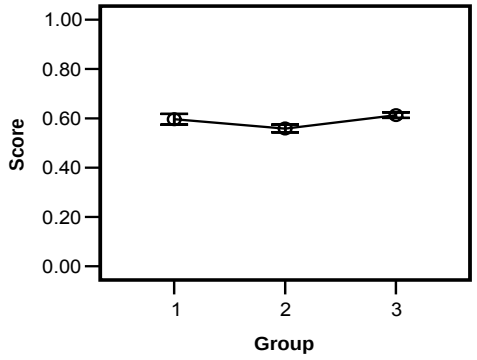
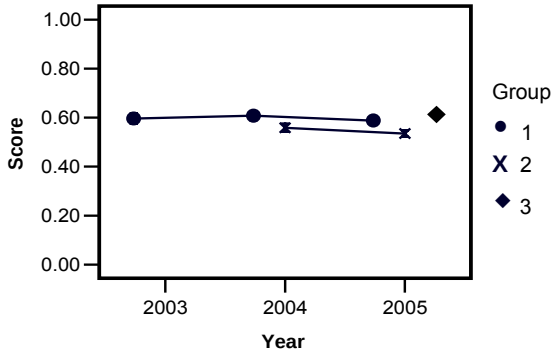


Exhibit 3.8
Students' Instructional Ratings Over Time



Personalized Relationships

We distilled two indexes to assess the extent to which relationships in NCHS schools were positive and supportive. One gauged student-to-student relationships and the other gauged student-to-teacher relationships. Select items from the student-to-student index are presented on the page to follow.

Personalized Relationships between students and teachers characterized by close continuous communication and each student having at least one adult to coordinate the support needed for the student to achieve postsecondary goals

Scale	Survey	Questions	Alpha
A. Student/student relationships	Student	13a, b, d; 14a, b, d-g	.71

Do you agree with the following statements…

Q14a Most students at this school care about each other
Q14b Most students in this school just look out for themselves
Q14d Serious fights often happen between students at this school (Reverse code)
Q14e There are groups or cliques of students who don't talk to other students (Reverse code)
Q14f Most students in this school would help each other if a problem came up
Q14g Most students in this school are mean to each other (Reverse code)

Exhibit 3.9 shows that students rated first-year relationships with other youth about the same across groups. Exhibit 3.10 shows that ratings in Group 2 schools dipped in 2004-05.

Exhibit 3.9
First-Year Youth Relationships

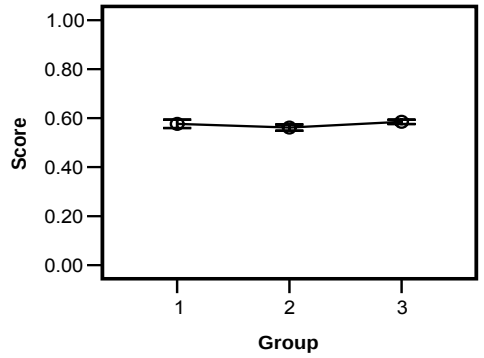
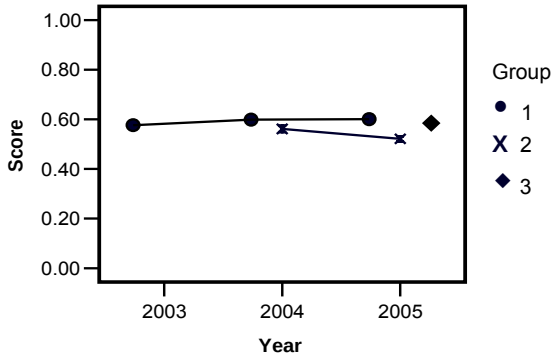


Exhibit 3.10
Youth Relationships Over Time



The second index assessing relationships measured student-to-teacher relationships, and select items are presented on the following page.

Personalized Relationships (continued)

Scale	Survey	Questions	Alpha
B. Student/teacher relationships	Student	13a, b, d; 15b, h; 17f; 18a, b, d, e; 19 a, c, e-h	.84

Do you agree with the following statements…

Q15b I feel that I can talk to the teacher in this school about things that are bothering me
Q15h I feel safe and comfortable with the teachers in this school

Do you talk with an adult in your school about the following at least once a week …

Q18a What's going on in your life
Q18b School or schoolwork
Q18d Your future goals and plans
Q18e College plans

Do you sometimes engage in the following activities during advisory period…

Q19a Talk about what's going on in your life or other students' lives outside of school
Q19c Get extra help with classes
Q19e Talk about colleges that you would like to attend
Q19f Talk about careers and future plans
Q19g Discuss current or world events
Q19h Learn or talk about skills you will need as an adult

Exhibit 3.11 shows that students rated the quality of first-year student-teacher relationships about the same across groups. Exhibit 3.12 shows that Group 2 ratings declined in 2004-05.

Exhibit 3.11
First-Year Teacher Relationships

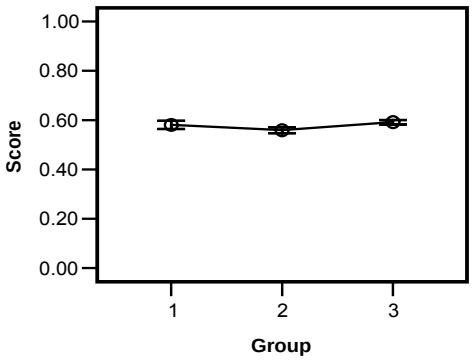
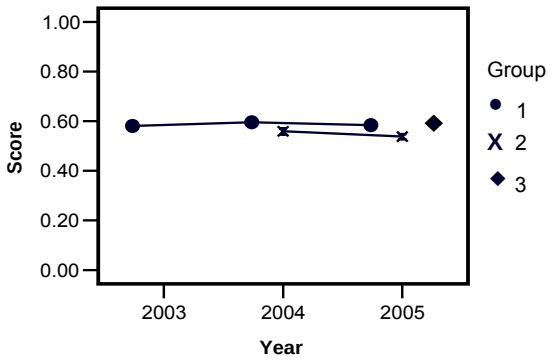


Exhibit 3.12
Teacher Relationships Over Time



Clear Instructional Focus

We developed one index to gauge instructional focus in NCHS schools. The index measured the alignment of expectations across teachers, and items are presented below.

Clear Focus on teaching and learning and expectations that every student will succeed

Scale	Survey	Questions	Alpha
A. Shared expectations	Teacher	11a, b, 12a, b, e, g, h, i	.67

Do you agree with the following statements…

Q11a Most teachers share the same beliefs and values about the central mission of the school
Q11b There is a great deal of cooperative effort among the staff members
Q12a I understand and support this school's educational focus
Q12b The school's educational focus is closely coordinated across grades
Q12e I use instructional strategies that are consistent with this school's educational focus
Q12g There is a common set of classroom assessments that I and all teachers use
Q12h There is a core curriculum that I and all teachers follow
Q12i My course content and instructional materials reflect the school's educational focus

Exhibit 3.13 shows that there were no significant differences by group in teachers' initial ratings of their schools' instructional focus. Exhibit 3.14 shows that teacher ratings started quite high and remained constant across school groups over time.

Exhibit 3.13
First-Year Instructional Focus

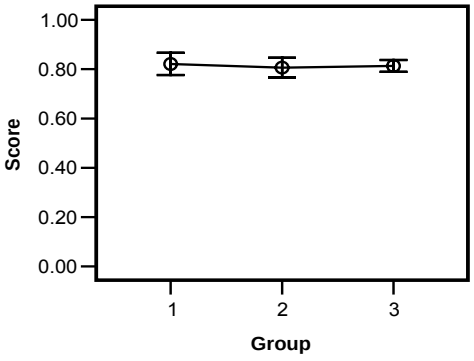
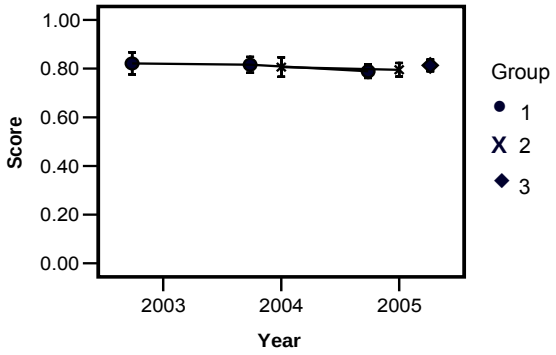


Exhibit 3.14
Instructional Focus Over Time



Instructional Leadership

We distilled one index to gauge the quality of the instructional leadership provided by principals. Survey items tapped teachers' perceptions of principals' support for and direct engagement with instruction. Select items are presented on the page to follow.

Instructional Leadership characterized by effective collaboration and school-wide support for teaching and learning

Scale	Survey	Questions	Alpha
A. Principal involved in instruction.	Teacher	35, 39, 41, 42	.90

Do you agree with the following…

Q35a Teachers in this school are evaluated fairly
Q35b Routine duties and paperwork interfere with teaching (reverse coded)
Q35c Teachers participate in making most of the important educational decisions in this school
Q35e Necessary instructional materials are available as needed by the staff
Q35f In this school, staff members are recognized for a job well done
Q35k There are formal arrangements within the school that provide opportunities for teachers to discuss and critique their own and others' instruction
Q39a My principal monitors the curriculum I use in my classroom, in part, to see that it reflects my school's educational focus
Q39b My principal monitors my classroom instructional practices, in part, to see that they reflect the school's educational focus
Q39c My principal evaluates my performance, in part, using criteria related to the school's educational focus
Q39d My principal praises, publicly recognizes, and/or provides tangible rewards to teachers whose instructional practices reflect the school's educational focus
Q39e My principal is available to provide me with guidance and assistance in structuring my instructional practices to reflect the school's educational focus
Q39f My principal informs teachers (in meetings, through written materials, etc.) about our progress in meeting our school goals

Does your principal visit your classroom for the following reasons weekly…

Q41b Observe your teaching
Q41b Model or demonstrate teaching strategies

Do you agree with the following statements regarding the feedback about teaching you receive from your principal…

Q42a It has been timely
Q42b It has focused on issues that I believe are important
Q42c It has focused on issues that I believe require my attention
Q42d It has included adequate follow-up and support for me to implement the recommended changes
Q42e It has been useful for improving my instruction
Q42f It has had an important effect on the quality of my teaching this year

Exhibit 3.15 shows that teachers' first-year ratings of the quality of principals' instructional leadership were higher in Group 2 and Group 3 schools than in Group 1 schools. Exhibit 3.16 shows that teacher ratings improved significantly in Group 1 schools over time but declined significantly in Group 2 schools—leaving all groups in about the same place on this measure by 2004-05.

Exhibit 3.15
First-Year Principal Leadership

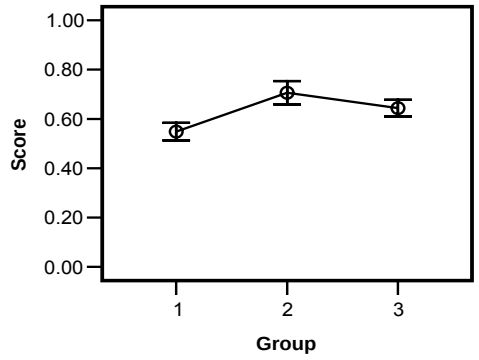
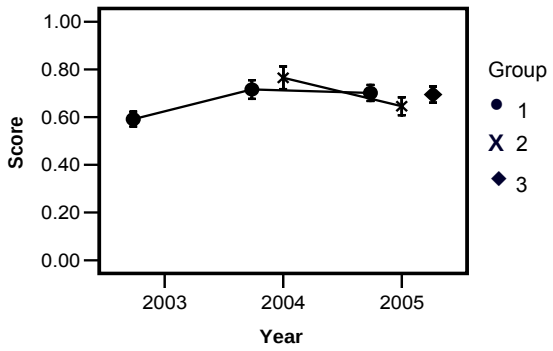


Exhibit 3.16
Principal Leadership Over Time



School-Based Teacher-Driven Professional Development and Collaboration

We distilled three indexes to assess professional development activities. One index measured teachers' perceptions of the amount of professional development they received. Another measured their perceptions of the quality of professional development. And the third index measured teachers' perceptions of their influence on school decision-making. The last index was included under this construct because participation in decision-making could be viewed as an informal type of professional development. Items from the index measuring the quantity of professional development are presented immediately below.

School-based Teacher-Driven Professional Development and Collaboration that is results-driven, standards-based, and embedded in the daily work of the school

Scale	Survey	Questions	Alpha
A. Amount of professional development	Teacher	36a-d, f-k;	.86

I received at least four hours of professional development this year in the following areas…

Q36a	Subject-specific content or instructional strategies
Q36b	Strategies for developing assessments
Q36c	Strategies for teaching low achieving students
Q36d	Methods for interpreting and using assessment data
Q36f	Methods for teaching literacy across content areas
Q36g	Methods for teaching inquiry skills across the curriculum
Q36h	Methods for thematic and interdisciplinary teaching
Q36i	Other general instructional strategies not listed above
Q36j	Professional development provided by or related to the community partner organization
Q36k	Professional development related to the school mission

Exhibit 3.17 shows that there were no statistically significant differences in teachers' first-year ratings of the quantity of their professional development opportunities. Exhibit 3.18 shows that teacher ratings remained consistent across groups over time.

Exhibit 3.17
First-Year Amount of PD

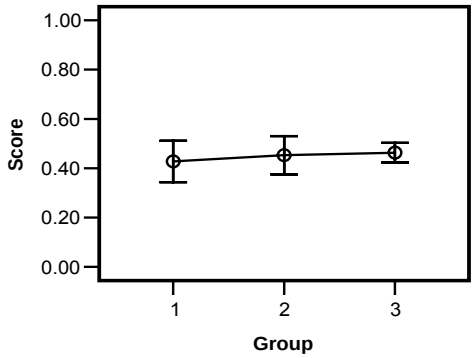
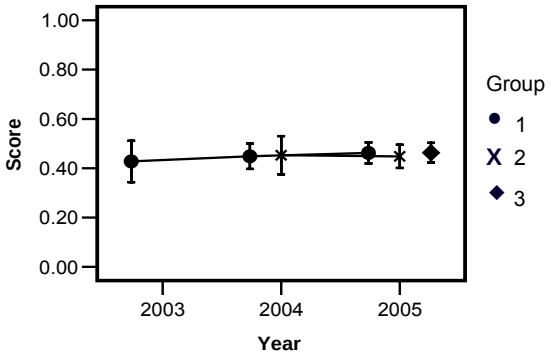


Exhibit 3.18
Amount of PD Over Time



The second index relating to professional development measured the *quality* of opportunities. It is presented immediately below.

School-based Teacher-Driven Professional Development and Collaboration (continued)

Scale	Survey	Questions	Alpha
B. Quality professional development	Teacher	37a-i, k-p; 38d, e	.94

The following statements usually applied to the professional development I received….

Q37a	Served my purposes
Q37b	Been a good use of my time
Q37c	Been appropriate to my own knowledge and skills
Q37d	Been appropriate for the grade level(s) or subject(s) I teach
Q37e	Been designed to communicate information about school, district, or state priorities, goals, or initiatives
Q37f	Helped me to reflect critically on how I teach and to develop plans for improving my teaching
Q37g	Given me new ideas and strategies to try in my classroom
Q37h	Included adequate follow-up or additional training to enable me to implement new ideas and strategies
Q37i	Included feedback and guidance while I was trying new strategies in my classroom
Q37k	Been sustained and coherently focused, rather than brief and unrelated
Q37l	Included opportunities to work productively with colleagues in my school
Q37m	Deepened my understanding of the subject matter I teach
Q37n	Led me to make changes in my teaching
Q37o	Helped my school staff work together better
Q37p	Deepened my understanding of how students learn
Q38d	Built effectiveness as a team (This question referred to planning meetings with colleagues.)

Exhibit 3.19 shows no statistically significant differences in teachers' first-year ratings of the quality of professional development. Exhibit 3.20 shows that teacher ratings of professional-development quality remained consistent across school groups over time.

Exhibit 3.19
First-Year Quality of PD

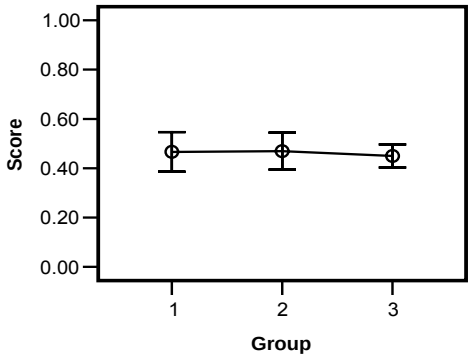
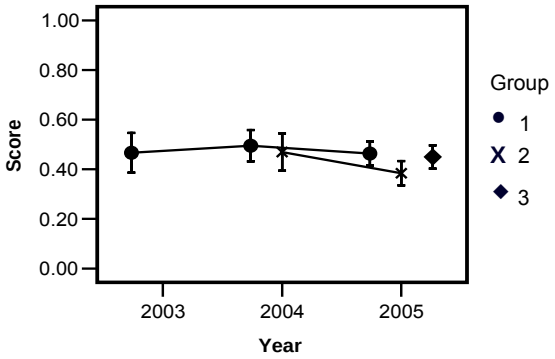


Exhibit 3.20
PD Quality Over Time



The third index relating to professional development measured the quality of teacher influence in school governance and related matters. It is presented immediately below.

School-based Teacher-Driven Professional Development and Collaboration (continued)			
Scale	**Survey**	**Questions**	**Alpha**
C. Teacher influence overall	Teacher	40a-j	.84

Did teachers have influence in the following areas…

Q40a	Determining the content and design of professional development for teachers
Q40b	Planning school budgets
Q40c	Establishing school discipline policies
Q40d	Establishing and shaping the school curriculum
Q40e	Determining the course schedule
Q40f	Establishing the school calendar
Q40g	Determining the school's goals and mission
Q40h	Selecting instructional materials that support the curriculum
Q40i	Determining student retention and promotion policy
Q40j	Making staffing and/or hiring decisions

Exhibit 3.21 shows that there were no statistically significant differences by school group in teachers' start-up-year ratings of the extent of their influence over school policy. Exhibit 3.22 shows that teacher ratings tended to decline over time. In Group 2 schools, the decline was statistically significant.

Exhibit 3.21
First-Year Teacher Influence

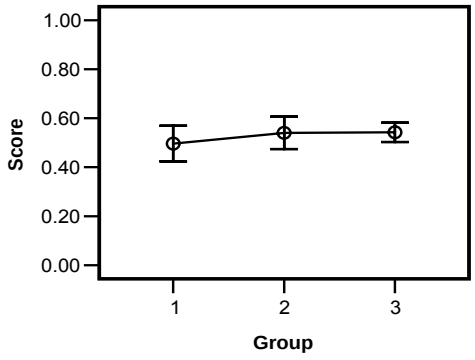
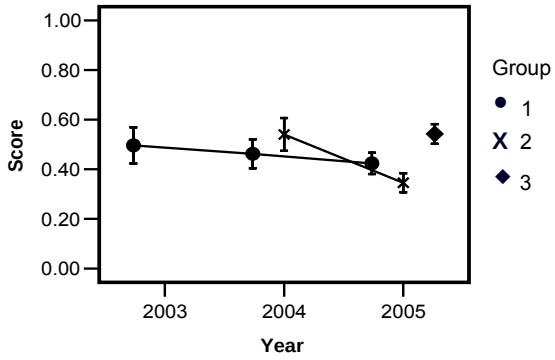


Exhibit 3.22
Teacher Influence Over Time



Meaningful Continuous Assessment

We developed an index to assess the quality of classroom assessments. The index measured students' perceptions of the appropriateness of classroom tests.

Meaningful Continuous Assessments to diagnose student needs and improve instruction			
Scale	**Survey**	**Questions**	**Alpha**
A. Tests are appropriate	Student	10 b, c, g	.68

Are the following statements true in some or all classes…

Q10b The tests are a good measure of how much I have learned
Q10c The tests cover the same material that the teacher presented in class
Q10g My teachers prepare me well for the tests

Exhibit 3.23 shows that Group 3 students provided very high initial ratings of classroom assessments. Exhibit 3.24 shows that student ratings were consistent over time.

Exhibit 3.23
First-Year Classroom Assessments

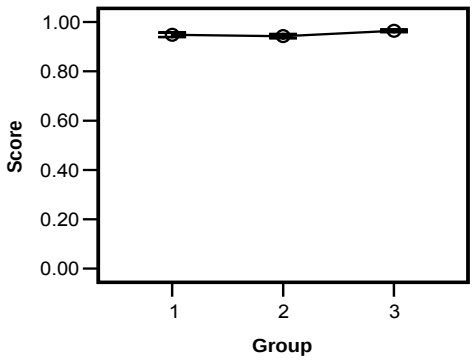
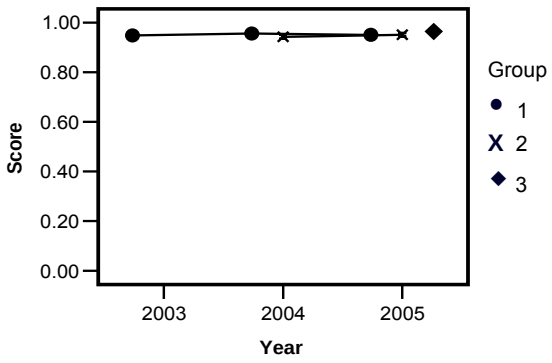


Exhibit 3.24
Classroom Assessments Over Time



Partnerships

We distilled two indexes to assess the support that partners provided to NCHS schools. One index measured partners' perceptions of their involvement in their NCHS schools. The other measured principals' perceptions of partners' involvement. The partner index is immediately below.

Partners offer support and opportunities for students

Scale	Survey	Questions	Alpha
Involvement in day-to-day operations, etc.	Partner	6b-m, o, q, r 15, 16e	.82

Do you have some role in the following…

Q6b Planning school budgets
Q6c Delivery of instruction
Q6d Tutoring and/or mentoring of students
Q6e Academic planning with students
Q6f Evaluating the overall instructional program
Q6g Teacher recruitment and hiring
Q6h Student recruitment and selection
Q6i Professional development planning with teachers
Q6j Administration
Q6k After-school programming
Q6l Curriculum design or selection
Q6m Determining the content and delivery of teacher professional development activities
Q6o Providing faculty professional development
Q6q Helping out in the building (e.g., as teachers' assistants, lunchroom monitoring)
Q6r Communicating with parents

Q15 Do most students at the school attend an event or volunteer at a program you organization sponsors weekly?

Q16e Did you sponsor career awareness activities that included most students at the school this year?

Exhibit 3.25 shows that there were no significant differences by group in partners' first-year ratings of involvement. Exhibit 3.26 shows that partner ratings remained consistent across school groups over time.

Exhibit 3.25
First-Year Partner Involvement

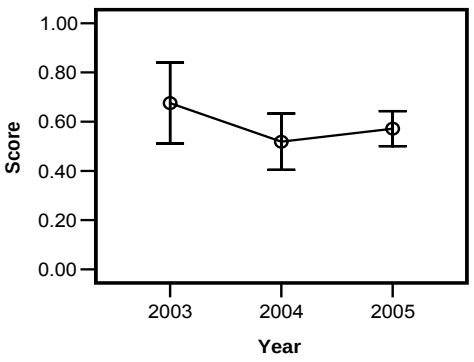
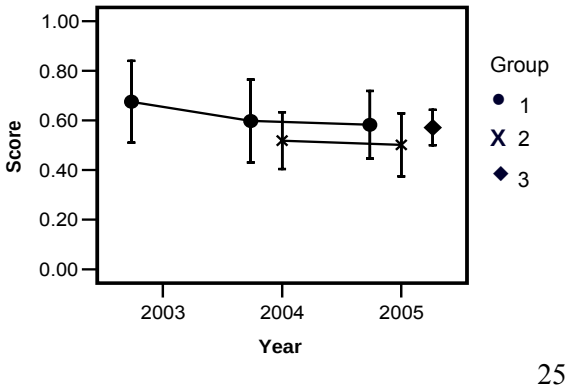


Exhibit 3.26
Partner Involvement Over Time



The second index measured principals' perceptions of community partners' engagement in school life. It is presented immediately below.

Partners (continued)

Scale	Survey	Questions	Alpha
Work with community partner	Principal	22a-r, 23, 24a-d	.83

Did your community partner have some role in the following areas of your school…

Q22a	Fund raising
Q22b	Planning school budgets
Q22c	Delivery of instruction
Q22d	Tutoring and/or mentoring students
Q22e	Academic planning with students
Q22f	Evaluating the overall instructional program
Q22g	Teacher recruitment and hiring
Q22h	Student recruitment and selection
Q22i	Professional development planning with teachers
Q22j	Administration
Q22k	After-school programming
Q22l	Curriculum design or selection
Q22m	Determining the content and delivery of teacher professional development activities
Q22n	Determining specific professional and teaching assignments
Q22o	Providing faculty professional development
Q22p	Organization of out-of-school learning opportunities for students
Q22q	Helping out in the building (e.g., as teachers' assistants, lunchroom monitoring)
Q22r	Communicating with parents

Exhibit 3.27 shows that there were no statistically significant differences by group in principals' first-year ratings of partners' involvement. Exhibit 3.28 shows that principals' ratings remained consistent within school groups over time. It is notable that principals' appraisal of partners' involvement was consistently lower than partners' self-appraisal.

Exhibit 3.27
First-Year Partner Involvement
(as reported by principal)

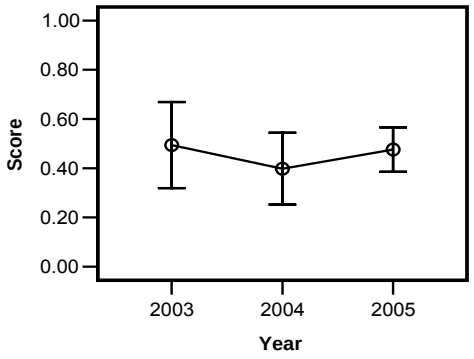
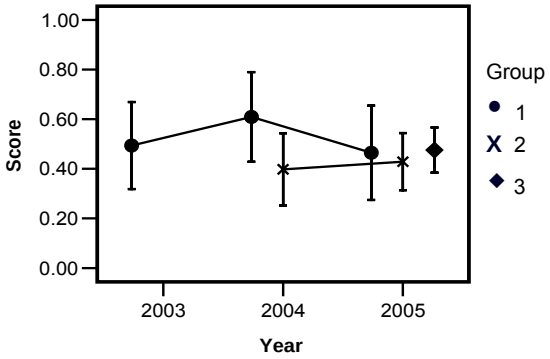


Exhibit 3.28
Partner Involvement Over Time
(as reported by principal)



Family/Caregiver Partnership and Involvement

We distilled two indexes to assess family involvement in NCHS schools. One index measured students' perceptions of their schools' outreach efforts. The other index measured students' perceptions of their families' level of engagement in school activities. The outreach scale is immediately below.

Family/Caregiver Partnership and Involvement in the governance and design of the school's education program includes two-way communication and extended learning opportunities

Scale	Survey	Questions	Alpha
Outreach to parents	Student	25a-c	.67

Does your parent, grandparent, or guardian …

Q25a Receive materials from the school that are mailed home
Q25b Receive invitations to events or meetings at the school
Q25c Have conversations with teachers or administrators over the phone

Exhibit 3.29 shows that Group 2 students rated their schools' first-year efforts in outreach to parents as more intensive than did students in other school groups. Exhibit 3.30 shows that ratings significantly declined in 2004-05 in both Group 1 and Group 2 schools.

Exhibit 3.29
First-Year Parent Outreach

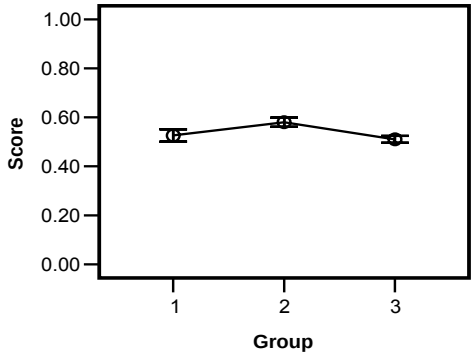
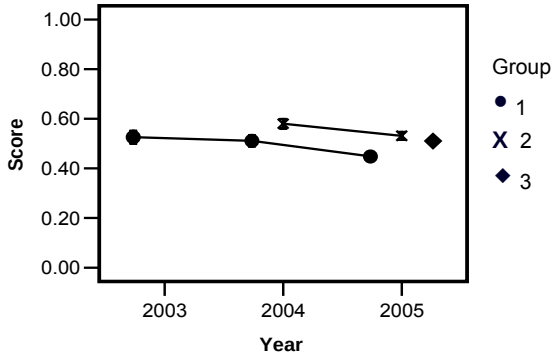


Exhibit 3.30
Parent Outreach Over Time



The second index measured students' perceptions of their caregivers' participation in school life, and it is presented on the page to follow.

Family/Caregiver Partnership and Involvement (continued)			
Scale	**Survey**	**Questions**	**Alpha**
Parent involvement	Student	25d-f	.80

Does your parent, grandparent, or guardian do the following…

Q25d Meet with teachers or administrators
Q25e Attend school events
Q25f Attend parent meetings hosted by the school

Exhibit 3.31 shows that there were no statistically significant differences by school group in students' first-year ratings of parent involvement. Exhibit 3.32 shows that student ratings of parent involvement declined in both Group 1 and Group 2 schools in 2004-05.

Exhibit 3.31
First-Year Parent Involvement

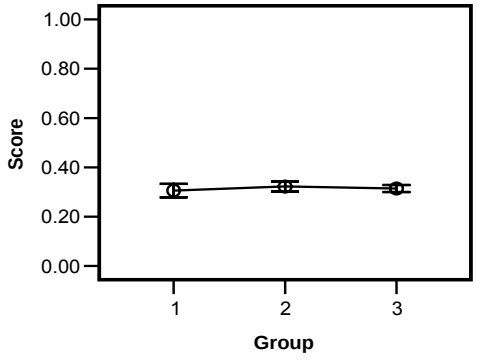
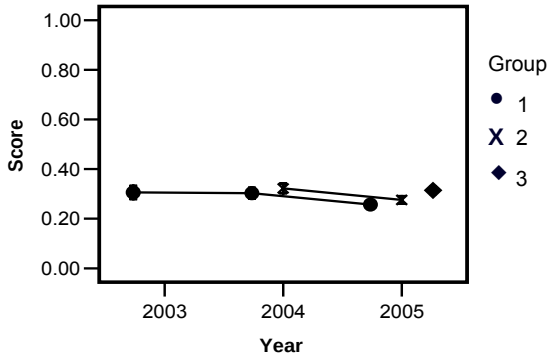


Exhibit 3.32
Parent Involvement Over Time



Youth Participation and Development

We distilled two indexes to assess youth participation and development in their schools. One index measured students' perceptions of advisories, since these sessions were designed, in large part, to facilitate student development. The other index measured student involvement in school leadership activities. The measure relating to advisories is presented on the page to follow.

Youth Participation and Development characterized by "student voice" in teaching and learning and shared responsibility for the operation and governance of the school

Scale	Survey	Questions	Alpha
Advisories	Student	19a,c-h	.84

Do you sometimes engage in the following activities in advisory periods…

Q19a Talk about what's going on in your life or other students' lives outside of school
Q19c Get extra help with classes
Q19d Talk about the college application process
Q19e Talk about colleges that you would like to attend
Q19f Talk about careers and future plans
Q19g Discuss current or world events
Q19h Learn or talk about skills you will need as an adult

Exhibit 3.33 shows that Group 2 students rated their schools' first-year advisories as providing fewer opportunities for the above-stated activities than did students in other groups. Exhibit 3.34 shows that student ratings remained consistent across groups over time.

Exhibit 3.33
First-Year Advisory Quality
(according to students)

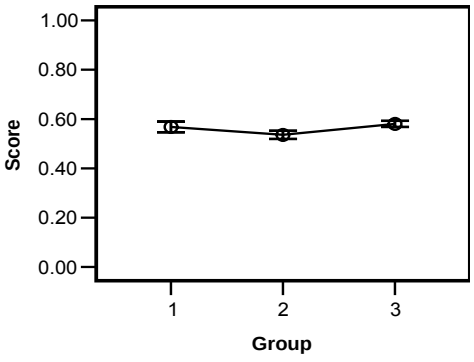
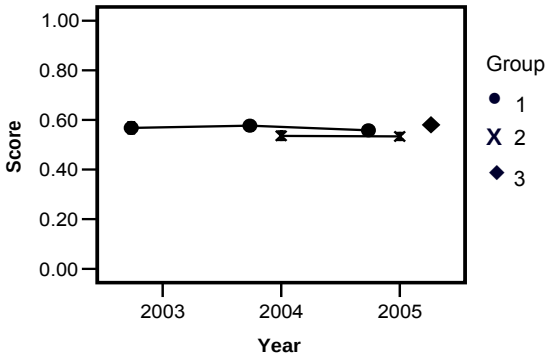


Exhibit 3.34
Advisory Quality Over Time
(according to students)



The second index measured student involvement in school leadership activities, and it is presented on the page to follow.

Youth Participation and Development (continued)

Scale	Survey	Questions	Alpha
Leadership activities	Student	16a-g, 24	.78

Did you participate in any of the following leadership opportunities at school…

Q16a Voted in student council elections
Q16b Volunteered or been selected to work in or lead an activity (sports, club, etc.)
Q16c Helped out in the school office (answered the phone, entered data in the computer, passed out in formation, etc.)
Q16d Served on a student council or leadership team for this school
Q16e Helped plan school events and activities
Q16f Helped with meetings for parents or community members
Q16g Been asked by staff for feedback/comments about the school or an activity

Q24 Did you volunteer during this school year to help out in the community through the school or organizations that work with the school (by tutoring another student, working in a soup kitchen, visiting the elderly, etc.)?

Exhibit 3.35 shows that students in Group 1 rated their first-year involvement in school and volunteer activities as more intensive than did students in other groups. Exhibit 3.36 shows that student ratings declined significantly in Group 1 schools in 2003-04 and 2004-05. Student ratings of involvement also declined significantly in 2004-05 in Group 2 schools.

Exhibit 3.35
First-Year Student Involvement in School and Volunteer Activities

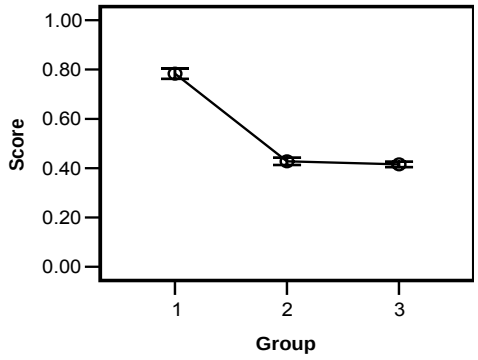
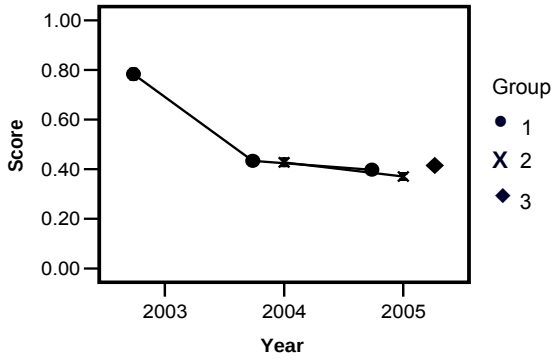


Exhibit 3.36
Student Involvement in School and Volunteer Activities Over Time



Effective Uses of Technology and Other Resources

We developed an index to assess the use of technology. That index measured teachers' instructional use of computers, and it is presented on the page to follow.

Effective Uses of Technology and Resources including print, visual, audio, and electronic			
Scale	**Survey**	**Questions**	**Alpha**
Use of computers	Teacher	34a, b	.69

Do you agree with the following statements…

Q34a In my classes, I teach students how to use computers through their course work (e.g., Internet research, word processing, spreadsheets, e-mail, etc.)

Q34b In my classes I give assignments that require students to use a computer

Exhibit 3.37 shows that in start-up years across groups teachers made comparable use of computers for instruction. Exhibit 3.38 shows that computer use by teachers was consistent within groups across years.

Exhibit 3.37
First-Year Computer Use

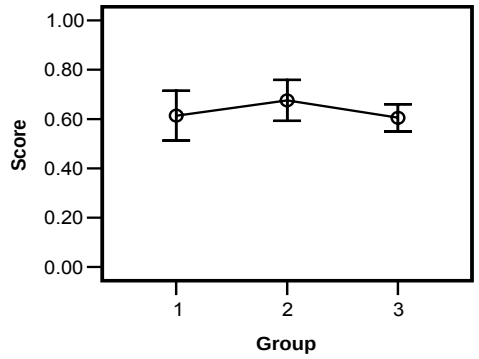
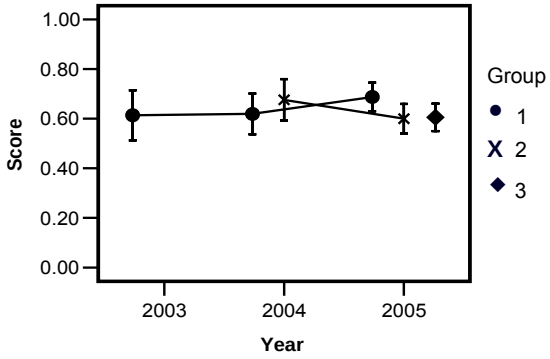


Exhibit 3.38
Computer Use Over Time



IV. FINDINGS REGARDING EDUCATIONAL OUTCOMES

The analyses presented in this chapter track the educational progress of students enrolled in NCHS schools. The data were obtained in electronic files from the New York City DOE. In this chapter, we first compare NCHS student achievement to citywide patterns. Next, we compare NCHS student achievement to the achievement of closely matched comparison-group students attending similar, but larger New York City public high schools.

Summary

We focused data analyses on three questions: What were the demographic characteristics of students who were enrolled in NCHS schools in 2004-05? Were these students headed for on-time graduation? Did the achievement of NCHS students exceed the achievement of comparison-group students? A brief summary, organized by research question, is presented immediately below, followed by a more detailed presentation of our findings and methods.

What Were the Demographic Characteristics of Students Enrolled in NCHS Schools in 2004-05?

Relative to city high school students, students in NCHS schools were more likely to be female, African American or Hispanic, and poor. They were less likely to be English Language Learners, new immigrants, or special education students. Lower proportions of NCHS students were proficient or advanced in English Language Arts (ELA) or math upon entry into high school than were proficient or advanced among eighth-graders citywide.

Were These Students Headed for On-time Graduation?

Using credit accumulation as the standard for high school graduation, 80 percent of ninth-grade entrants to NCHS schools were categorized as "on track" for graduation after one year, 68 percent after two years, and 77 percent after three years. While some NCHS schools has a practice of postponing Regents until later high school years, on average, as of 2004-05, NCHS students were passing Regents exams at a rate just about adequate for on-time graduation. By the end of tenth grade, the average NCHS student had passed 2.08 of the five Regents exams required for graduation. By the end of eleventh grade, the average NCHS student had passed 3.49 Regents exams.

Did the Achievement of NCHS Students Exceed the Achievement of Comparison-Group Students?

Rates of school attendance, credit accumulation, and promotion were higher among NCHS students than among comparison-group students, and the differences were statistically significant. NCHS schools also appeared to have better holding power than comparison schools, although the differences were not statistically significant. With regard to passing Regents exams, ninth-grade comparison-group students out-performed the first group of NCHS schools (those initiated in 2002-03) but under-performed the second group (those initiated in 2003-04). Suspension rates were higher among NCHS students in 2004-05 than among comparison-group students.

Findings from Descriptive Analyses of NCHS Students

Characteristics of Students at Enrollment in NCHS Schools

Exhibit 4-1 presents descriptive statistics on students enrolled for at least one day in NCHS schools in 2004-05. These statistics are presented alongside data on all New York City students enrolled in 2003-04, the most recent year for which aggregate data were available. Citywide data were obtained from DOE's 2003-04 publicly available school report cards and DOE's 2003-04 Report Card database. Because NCHS students were included in citywide aggregates, the following discussion is based on inspection of NCHS and city datasets rather than on statistical tests.

Exhibit 4-1 shows that more than half of NCHS students were in the ninth grade in 2004-05. This was disproportionately high but expected, given that most NCHS schools were still in start-up mode. Demographically, the schools enrolled more females (53 percent) than males (47 percent). Most students were African American or Hispanic (92 percent) and eligible for free and reduced-price lunch (81 percent). Lower proportions were English Language Learners (11 percent), new immigrants (8 percent), or eligible for special education (8 percent). In the eighth grade, about 26 percent of NCHS students were proficient or advanced in ELA, and 30 percent were proficient or advanced in math.

Relative to patterns in public high schools citywide, NCHS schools had higher proportions of ninth-grade students, females, African Americans, Hispanics, and children in poverty but lower proportions of English Language Learners, new immigrants and special education students. NCHS schools enrolled lower proportions of students identified in the eighth grade as proficient or advanced in ELA or proficient or advanced in math.

Exhibit 4.1
Characteristics of NCHS and New York City High School Students, in Percents

Characteristics	NCHS 2004-05	NYC 2003-04
All Grades	*(N = 12,111)*	*(N = 285,000)*[a]
Grade 9	56.5	34.6 [a]
Grade 10	31.5	28.3 [a]
Grade 11	8.6	16.6 [a]
Grade 12	2.2	14.4 [a]
Ungraded	1.2	6.1 [a]
Gender	*(N = 12,111)*	
Female	52.9	49.6 [b]
Male	47.1	50.4 [b]
Race	*(N = 11,903)*	
African American	41.9	35.0 [b]
Asian and Others	5.1	14.1 [b]
Hispanic	49.6	35.7 [b]
White	3.4	15.2 [b]
Eligible for ELL	*(N = 12,111)*	
Yes	10.9	12.9 [b]
Eligible for Free/Reduced Price Lunch	*(N = 12,111)*	
Eligible	81.0	53.9 [b]
Ineligible	10.0	
Missing Data	9.0	
Recent Immigrant	*(N =12,111)*	
Yes	7.7	11.4 [b]
Special Education	*(N = 12,111)*	
Yes	7.6[c]	10.6 [bc]
Eighth-Grade ELA Proficiency	*(N = 9,080)*	
Proficient/Advanced	25.6[d]	32.3 [b d]
Eighth-Grade Math Proficiency	*(N = 9,663)*	
Proficient/Advanced	29.9 [d]	34.1 [b d]

[a] Data for total enrollment disaggregated by grade were obtained from 2003-04 DOE School Report Card Database.

[b] Citywide percentages (with the exception of data disaggregated by grade) appear as published in 2003-04 DOE School Report Cards available at: http://www.nycenet.edu/daa/SchoolReports/ . These report cards are generated by the Report Card Database, which provides two separate variables for enrollment: 1) enrollment by school; and 2) enrollment by grade within school. In the database, total aggregated enrollment was 285,046 when aggregating at the school level and 285,000 when aggregating first by grade, then by school. It is not known to which total these percents refer, though they are extremely close. There were two schools in the database for which there was no enrollment data at all, and another 16 schools for which the grade-level enrollment data (when aggregated at the school level) was slightly off from the school-level total (which was a separate variable).

[c] Citywide special education percent includes part- and full-time special education students, as does the NCHS percentage.

[d] NCHS proficiency rates are reported for students in all grades. Citywide proficiency rates are reported for entering ninth- and tenth-graders, as presented in the School Report cards available online.

Educational Status of NCHS Students in 2004-05

We examined the educational status of NCHS students in five domains: school attendance, credit accumulation, grade promotion, passing of Regents exams, and suspension rates. Findings are presented serially.

Exhibit 4.2 presents the **average daily attendance** of students enrolled for at least one day in an NCHS school in 2004-05. The data are partitioned by students' grade, race, and gender. Across subgroups, the average daily attendance rate was 86 percent; the median was 92 percent. These statistics compared favorably to the average rate of attendance for high school students citywide, which was 83 percent in 2003-04 and 81 percent in 2004-05. We found no consistent differences in the attendance of NCHS students by grade, although attendance rates were highest among students in the eleventh grade and lowest among students in the twelfth grade. The attendance of boys and girls was about the same across grades. Attendance was highest among Asians and lowest among Hispanics. It was substantially lower in un-graded special education classes than in the mainstream.

Exhibit 4.2
NCHS Student Attendance Rates, 2004-05

	Overall		Ninth		Tenth		Eleventh		Twelfth		Ungraded Special Education	
	Percent	SD	Percent (*n*)	SD	Percent (*n*)	SD	Percent (*n*)	SD	Percent (*n*)	SD	Percent (*n*)	SD
African-American												
Male (n=2299)	**86.0**	**19.6**	85.7 (n=1409)	19.8	86.4 (n=662)	19.7	91.3 (n=151)	13.9	87.3 (n=22)	12.3	76.3 (n=55)	24.5
Female (n=2694)	**86.7**	**17.9**	87.3 (n=1575)	18.0	85.3 (n=851)	17.9	88.5 (n=204)	13.6	88.6 (n=40)	9.8	69.1 (n=24)	31.7
Total (n=4993)	**86.4**	**18.7**	86.5 (n=2984)	18.9	85.8 (n=1513)	18.7	89.7 (n=355)	13.8	88.1 (n=62)	10.7	74.1 (n=79)	26.9
Asian												
Male (n=230)	**92.6**	**13.5**	92.0 (n=123)	15.1	92.7 (n=71)	12.7	95.5 (n=34)	6.5	74.0 (n=2)	21.6		
Female (n=230)	**94.4**	**9.6**	95.2 (n=115)	6.5	92.2 (n=74)	14.2	96.6 (n=39)	4.1	83.2 (n=2)	6.8		
Total (n=460)	**93.5**	**11.7**	93.6 (n=238)	11.8	92.4 (n=145)	13.4	96.1 (n=73)	5.4	78.6 (n=4)	14.1		
Hispanic												
Male (n=2807)	**84.7**	**18.8**	85.1 (n=1510)	18.7	83.5 (n=922)	19.5	88.9 (n=259)	13.4	87.0 (n=69)	12.7	69.6 (n=47)	27.3
Female (n=3103)	**83.5**	**19.0**	84.5 (n=1636)	19.1	82.1 (n=1026)	19.9	84.8 (n=296)	16.2	81.3 (n=121)	15.0	74.8 (n=24)	15.2
Total (n=5910)	**84.1**	**18.9**	84.8 (n=3146)	18.9	82.8 (n=1948)	19.7	86.7 (n=555)	15.1	83.4 (n=190)	14.4	71.4 (n=71)	23.9
White												
Male (n=230)	**88.9**	**15.2**	87.2 (n=127)	17.5	89.7 (n=72)	12.5	95.2 (n=28)	4.4	81.2 (n=3)	24.0		
Female (n=169)	**85.6**	**20.8**	86.3 (n=96)	21.8	85.8 (n=45)	17.9	86.9 (n=20)	21.8	71.4 (n=7)	22.2	88.4 (n=1)	n/a
Total (n=399)	**87.5**	**17.8**	86.8 (n=223)	19.4	88.2 (n=117)	14.9	91.7 (n=48)	91.7	74.3 (n=10)	21.9	88.4 (n=1)	n/a
Other												
Male (n=66)	**83.2**	**21.5**	78.8 (n=34)	22.6	87.9 (n=32)	19.5						
Female (n=80)	**88.8**	**14.8**	90.9 (n=37)	9.7	87.3 (n=38)	18.8	84.1 (n=5)	11.9				
Total (n=146)	**86.3**	**18.3**	85.1 (n=71)	18.1	87.6 (n=70)	19.0	84.1 (n=5)	11.9				
Male (n=5709)	**85.8**	**18.9**	85.8 (n=3272)	19.1	85.3 (n=1767)	19.2	90.5 (n=472)	13.0	86.6 (n=96)	13.0	73.2 (n=102)	25.9
Female (n=6402)	**85.6**	**18.3**	86.5 (n=3568)	18.2	84.1 (n=2049)	18.9	87.1 (n=566)	15.2	82.6 (n=170)	14.6	72.3 (n=49)	24.7
Total (n=12,111)	**85.7**	**18.6**	86.2 (n=6840)	18.7	84.7 (n=3816)	19.1	88.6 (n=1038)	14.4	84.1 (n=266)	14.2	72.9 (n=151)	25.5

Exhibit reads: African-American males attended NCHS schools 86 percent of the days school was in session in 2004-05.

We examined **credit accumulation** among students who entered NCHS schools as ninth-graders and were still enrolled in 2004-05. This group included ninth-grade entrants from 2004-05 with one year of data, ninth-grade entrants from 2003-04 with

two years of data, and ninth-grade entrants from 2002-03 with three years of data. Findings by the years students were in an NCHS high school are as follows:

- Among the 4,776 NCHS students enrolled as ninth-graders in 2004-05 for whom we have credit data, credit accumulation averaged 10.47 units (the median was 11.33 units), which indicates that the average NCHS student was on track for graduation after one year in an NCHS school. To be counted as making timely progress toward graduation, ninth-grade students in New York City schools needed to accumulate at least eight units of course credit.

- Among the 2,404 NCHS students enrolled as ninth-graders in 2003-04 for whom we have two years of credit data, credit accumulation averaged 21.62 units (the median was 23.25 units), meaning that the average NCHS student was on track for graduation after two years. To be counted as making timely progress toward graduation, tenth-grade students needed to accumulate at least 20 units of credit.

- Among the 715 NCHS students enrolled as ninth-graders in 2002-03 for whom we have three years of credit data, credit accumulation averaged 33.62 units (the median was 35.35 units), meaning that the average NCHS student was also on track for graduation after three years. To be counted as making timely progress toward graduation, eleventh-grade students in New York City schools in 2004-05 needed to accumulate at least 28 units of credit.

- Because NCHS schools had been in operation for a maximum of three years in 2004-05, no student had four years of exposure to an NCHS school. On average, however, NCHS students were accumulating more than adequate numbers of credits for on-time graduation at each earlier grade level (10-11 units of credit annually), so the prognosis for on-time graduation among twelfth-grade students in 2005-06 was good based on credit accumulation.

Using credit accumulation as the sole measure of the likelihood of high school graduation, 80 percent of ninth-grade entrants to NCHS schools were categorized as "on track" for graduation after one year, 68 percent after two years, and 77 percent after three years. These rates are promising in light of citywide graduation rates. According to DOE figures, 54 percent of students graduated on time in 2004, up from 51 percent in 2001. For black and Hispanic students on-time graduation rates were lower. Only 49 percent of black and Hispanic students graduated on time in 2004.

We also examined rates of **grade promotion** among NCHS students. Of the 6,225 ninth-grade students enrolled in NCHS schools in 2004-05 for one day or more and still active somewhere in the New York City school system in 2005-06, 84 percent were promoted, as indicated by their 2005-06 grade status. The citywide ninth-grade

promotion rate in 2004-05 was 72 percent. The comparable numbers for NCHS tenth- and eleventh-graders were 80 percent (of 3,186 students) and 93 percent (of 958 students), respectively. As the on-track credit findings above indicate, the move from tenth to eleventh grade proved to be more difficult for NCHS students than that from ninth to tenth or eleventh to twelfth grade. For the small number of seniors in NCHS schools in 2004-05 (numbering 234 students), all of whom had enrolled after ninth grade, graduation rates in 2004-05 (based on their grade and discharge codes the following year) were relatively high at 87 percent.

To earn a high school diploma, in addition to passing courses, New York City students must pass state-level subject-area tests known as **Regents examinations**. Students are not required to pass a certain number of Regents tests at each grade, but they must pass a certain number Regents exams before graduation. Beginning with the class of 2005, students who scored at least 55 on each of five state exams received a local diploma, and students who scored at least 65 on each of five state exams received a Regents diploma. Advanced Regents diplomas were awarded to students who scored 65 or better on eight exams. In 2004, only 18 percent of New York City students received Regents diplomas.

Exhibit 4.3 shows the numbers of Regents tests passed by NCHS students with a score of 55 or higher as of 2004-05. By the end of tenth grade, the average NCHS student had passed 2.08 (median 2.00) of the five Regents exams required for graduation. By the end of eleventh grade, the average NCHS student had passed 3.49 (median 4.00) Regents exams. Despite a policy in some NCHS schools of postponing Regents testing, the data indicate that NCHS students were, on average, passing just about enough Regents exams to graduate on time.

NCHS students were not immune to trouble, as evidenced in **suspension** rates. Of the 12,111 students enrolled in 2004-05 in an NCHS school for one day or more, 710 were suspended, or nearly 6 percent, which was roughly comparable to the citywide suspension rate of 6 percent in 2003-04.

Suspension rates were substantially higher in NCHS schools situated on Impact School campuses. Among the 8,114 NCHS students attending a school on an Impact campus, 8 percent were suspended. Among the 3,997 NCHS students attending a school on a non-Impact campus, 5 percent were suspended.

Among general education students in NCHS schools, suspension rates were highest for ninth-grade and eleventh-grade students (6 percent), with tenth-grade students slightly lower (5 percent). Suspension rates in general education were lowest for twelfth-grade students (3 percent). About 8 percent of NCHS students in ungraded special education were suspended in 2004-05.

In NCHS schools, males were more likely to be suspended (7 percent) than females (5 percent). African Americans (8 percent), “Others” (8 percent), and

Hispanics (5 percent) were more likely to be suspended than whites (2 percent) and Asians (1 percent).

Exhibit 4.3
Number of Regents Tests Passed by NCHS Students, 2004-05

	Overall		Ninth		Tenth		Eleventh		Twelfth		Ungraded Special Education	
	Avg. Passed	SD	Avg. Passed (*n*)	SD	Avg. Passed (*n*)	SD	Avg. Passed (*n*)	SD	Avg. Passed (*n*)	SD	Avg. Passed (*n*	SD
African-American												
Male (n=2299)	**1.10**	**1.53**	0.43 (n=1409)	0.76	2.01 (n=662)	1.53	3.34 (n=151)	2.28	3.77 (n=22)	2.02	0.11 (n=55)	0.69
Female (n=2694)	**1.26**	**1.58**	0.46 (n=1575)	0.70	2.15 (n=851)	1.53	3.30 (n=204)	2.13	4.30 (n=40)	1.92	0.17 (n=24)	0.64
Total (n=4993)	**1.19**	**1.56**	0.45 (n=2984)	0.73	2.09 (n=1513)	1.53	3.32 (n=355)	2.19	4.11 (n=62)	1.96	0.13 (n=79)	0.67
Asian												
Male (n=230)	**1.33**	**1.85**	0.29 (n=123)	0.73	1.77 (n=71)	1.73	3.97 (n=34)	1.80	4.50 (n=2)	0.71		
Female (n=230)	**1.41**	**1.81**	0.47 (n=115)	0.76	1.53 (n=74)	1.71	3.69 (n=39)	1.72	6.50 (n=2)	0.71		
Total (n=460)	**1.37**	**1.83**	0.38 (n=238)	0.75	1.65 (n=145)	1.72	3.82 (n=73)	1.75	5.50 (n=4)	1.29		
Hispanic												
Male (n=2807)	**1.37**	**1.80**	0.45 (n=1510)	0.81	2.08 (n=922)	1.66	3.47 (n=259)	2.35	5.10 (n=69)	1.31	0.26 (n=47)	1.01
Female (n=3103)	**1.45**	**1.87**	0.42 (n=1636)	0.73	2.12 (n=1026)	1.74	3.54 (n=296)	2.37	4.84 (n=121)	1.61	0.08 (n=24)	0.41
Total (n=5910)	**1.41**	**1.84**	0.43 (n=3146)	0.77	2.10 (n=1948)	1.70	3.51 (n=555)	2.36	4.94 (n=190)	1.51	0.20 (n=71)	0.86
White												
Male (n=230)	**1.55**	**2.01**	0.53 (n=127)	0.96	2.10 (n=72)	1.80	4.46 (n=28)	2.17	4.33 (n=3)	4.04		
Female (n=169)	**1.40**	**1.96**	0.35 (n=96)	0.65	2.11 (n=45)	1.93	3.90 (n=20)	2.51	4.29 (n=7)	1.25	0.00 (n=1)	n/a
Total (n=399)	**1.49**	**1.99**	0.45 (n=223)	0.84	2.10 (n=117)	1.84	4.23 (n=48)	2.30	4.30 (n=10)	2.16	0.00 (n=1)	n/a
Other												
Male (n=66)	**1.23**	**1.49**	0.41 (n=34)	0.66	2.09 (n=32)	1.63						
Female (n=80)	**1.44**	**1.35**	0.54 (n=37)	0.56	2.05 (n=38)	1.27	3.40 (n=5)	1.52				
Total (n=146)	**1.34**	**1.41**	0.48 (n=71)	0.61	2.07 (n=70)	1.44	3.40 (n=5)	1.52				
Male (n=5709)	**1.26**	**1.70**	0.43 (n=3272)	0.79	2.04 (n=1767)	1.62	3.53 (n=472)	2.29	4.76 (n=96)	1.67	0.18 (n=102)	0.85
Female (n=6402)	**1.35**	**1.74**	0.44 (n=3568)	0.71	2.11 (n=2049)	1.65	3.47 (n=566)	2.24	4.71 (n=170)	1.69	0.12 (n=49)	0.53
Total (n=12,111)	**1.31**	**1.72**	0.44 (n=6840)	0.75	2.08 (n=3816)	1.64	3.49 (n=1038)	2.27	4.73 (n=266)	1.68	0.16 (n=151)	0.76

Exhibit reads: Across grades 9-12, African-American males in NCHS schools passed an average of 1.1 Regents exams by 2004-05.

Suspension rates were lower in NCHS schools before 2004-05. Of the 1,169 students enrolled in 2002-03, only 2 percent were suspended that year. Of the 4,258 students enrolled in 2003-04, only 4 percent were suspended in that year.

Findings from Matched Comparison Study

In the preceding section, we examined the achievement of NCHS students against citywide norms. Below we present analyses comparing NCHS students with comparable students in demographically similar but larger New York City public high schools. We address six outcomes: holding power, attendance rates, credit accumulation, promotion rates, Regents exams passed, and suspension rates. Analyses are disaggregated, when appropriate, by school group and student cohort. School groups are defined in terms of the year that each set of NCHS schools opened: Group 1 schools opened in 2002-03, Group 2 schools opened in 2003-04, and Group 3 schools opened in 2004-05. Cohorts are defined by students' grade and year of entry in an NCHS school: for example, Cohort 1 students entered an NCHS school as ninth-graders in 2002-03, Cohort 2 students entered an NCHS school as tenth-graders in

2002-03, Cohort 3 students entered an NCHS school as eleventh-graders in 2002-03, Cohort 4 students entered an NCHS school as twelfth-graders in 2002-03, Cohort 5 students entered an NCHS school as ninth-graders in 2003-04, etc.

We selected comparison-group students for these analyses through a multi-step process that employed propensity-score matching. The process is explained fully in the methods chapter and appendix. First, we matched the 30 NCHS schools operating in 2003-04 with 12 larger traditional public high schools (each with a capacity equal to or greater than 750 students). Matches were selected from among all New York city high schools based on: the percent of their students who were eligible for free and reduced price lunch; the percent who were recent immigrants; the percent who were English Language Learners; the percent African American, Hispanic, Asian, White, and Other; and students' eighth-grade average English Language Arts and math scale scores on statewide tests. Having selected a pool of comparison schools, we next found matches for NCHS students within the pool of schools. We specifically matched ninth-grade entrants to NCHS schools to ninth-grade entrants in comparison-group schools, based on: students' eighth-grade academic records (including attendance, math scale score, and English Language Arts scale score) and background characteristics (including gender, race, age, immigrant status, special education eligibility, ELL eligibility, and free or reduced-price lunch eligibility.)

Each analysis employed matched pairs of students who entered high school in the ninth grade and for whom we had data for the outcome in question. Because we were interested in the effect of NCHS schools on treated students, research samples—other than those used in the holding power analyses—were further constrained to include only persisters, whom we defined as students who were enrolled in ninth grade in a sample school for 150 days and were not discharged in the year or years covered by the analysis. Analyses of Regents test performance and suspension data both employed this constraint. For attendance and credit accumulation analyses, research samples were additionally constrained to include only students with data for both variables. Finally, to be included in promotion analyses, students were required to have data in their records for attendance, credit accumulation, and promotion.

The treatment of attendance and credit accumulation is especially detailed in the discussion that follows because those data could meaningfully be considered for both one-year persisters and multi-year persisters. Other analyses are less expansive due either to the nature of the measurement opportunity (e.g., there were no grade-level Regents test requirements) or the rarity of the event (suspension).

To summarize briefly across indicators, available outcome data for paired comparisons show that NCHS attendance, credit accumulation, and promotion rates exceeded comparison-group rates, and differences were statistically significant. Holding power was better in NCHS schools, although differences were not statistically significant, and Regents test performance was better in Group 2 NCHS schools than in the schools attended by comparison students.

Findings favoring NCHS schools were reversed only twice. Suspension rates in 2004-05 were lower among comparison-group students. Also, for students admitted to ninth grade in 2002-03, Regents performance in 2004-05 was higher among comparison-group students.

Results from these outcomes analyses were anticipated and validated by the implementation data. As predicted by implementation results, outcomes improved for entrants after 2002-03 and were best for entrants to NCHS schools launched in 2003-04.

We begin our detailed discussion of outcomes for paired comparisons with an examination of students' **persistence** rates, with persistence defined as a student's continued enrollment in the school that he or she entered in ninth grade. (More specifically, persisters were defined as students enrolled for at least 150 days in a sample school and not "discharged." An enrollment standard was necessary because discharge data, used alone, were found to be unreliable.) Non-persisters included both transfers and dropouts. We clustered these groups together because available data did not enable us to distinguish reliably between the two.

Our analyses examined 220 student pairs who enrolled in ninth grade in 2002-03 (Cohort 1). These students were expected to begin their third year in a sample high school in 2004-05. Among NCHS students, 183 students or 83 percent persisted until 2004-05. Among comparison-group students 169 students or 77 percent persisted over the same period. Fourteen more NCHS students than comparison-group students persisted into the third year—a 6 percent difference, which is notable but not statistically significant.

Analyses also examined 652 student pairs who enrolled in ninth grade in the sample schools in 2003-04 (Cohort 5). These students were expected to begin their second year in a sample school in 2004-05. Among NCHS students, 584 students or 90 percent re-entered a sample school in 2004-05. Among comparison-group students, 568 students or 87 percent re-entered a sample school in 2004-05. Sixteen more NCHS students than comparison-group students persisted into the second year—a 3 percent difference, which is again notable but not statistically significant.

We examined **attendance and credit accumulation** data from two vantage points. To align outcome analyses with implementation analyses, we compared the performance of ninth-grade students in the start-up year of each NCHS school group. To understand longitudinal effects of the NCHS schools, we compared two-year outcomes for persisting ninth-graders admitted in 2003-04, and three-year outcomes for persisting ninth-graders admitted in 2002-03. Persisters were defined as ninth-grade students who were enrolled for 150 days and not discharged.

The exhibits that follow present data in three panels. The top panel identifies the number of student pairs in the analysis by cohort and school group. The middle panel describes the performance of NCHS students. The bottom panel describes the

performance of comparison students. Where the difference between NCHS and comparison students is statistically significant, that fact is indicated with an asterisk in the cell of the higher performing group, and the effect size is shown in parentheses. Larger effect sizes indicate larger differences.

Exhibit 4.4 presents **attendance** patterns for one-year persisters. Ninth-grade attendance was significantly better among NCHS students in four of six group/cohort combinations. For students admitted in 2002-03 and for Cohort 9 Group 3 students admitted in 2004-05, the attendance of NCHS and comparison students was equivalent. Effect sizes were generally highest for students admitted in 2003-04, indicating that the difference between NCHS and comparison students was greatest for those students.

Exhibit 4.4
For Paired Comparisons, One-Year Attendance Rates

Sample Size (n=1630)

	C1	C5	C9
Group 1 (schools starting in 2002-03)	169	146	138
Group 2 (schools starting in 2003-04)		426	275
Group 3 (schools starting in 2004-05)			476

NCHS Average Attendance Rates (in percents)

	C1	C5	C9
Group 1	92.26	93.08* (0.26)	92.85* (0.21)
Group 2		92.65* (0.19)	92.78* (0.15)
Group 3			91.94

Comparison Group Average Attendance Rates (in percents)

	C1	C5	C9
Group 1	93.07	90.53	90.38
Group 2		90.16	91.09
Group 3			91.11

Exhibit reads: NCHS ninth-graders in Cohort 1 (n=169) had an average attendance rate of 92.26 percent, which was not significantly different from the attendance rate of their matches in comparison schools, which was 93.07 percent.

* Statistically significant at $p \leq .05$. Effect sizes are shown in parentheses.

Exhibit 4.5 presents longitudinal attendance patterns, which mirrored ninth-grade outcomes. As was true for ninth-graders in their ninth-grade year, NCHS attendance was significantly better than comparison-group attendance for two-year persisters (members of Cohort 5). NCHS and comparison group attendance was equivalent for three-year persisters (members of Cohort 1).

Exhibit 4.5
For Paired Comparisons, Two-Year and Three-Year Attendance Rates

Sample Size (n=526)		
	C1 (3 years)	C5 (2 years)
Group 1 (schools starting in 2002-03)	105	109
Group 2 (schools starting in 2003-04)		312
Group 3 (schools starting in 2004-05)		

NCHS Average Attendance Rates (in percents)		
	C1 (3 years)	C5 (2 years)
Group 1	91.86	92.34* (0.25)
Group 2		92.68* (0.19)
Group 3		

Comparison Group Average Attendance Rates (in percents)		
	C1 (3 years)	C5 (2 years)
Group 1	91.47	89.68
Group 2		90.41
Group 3		

Exhibit reads: NCHS students in Cohort 1 (n=105) had an average attendance rate of 91.86 percent over three years. This was not significantly different from the attendance rate of their matches in comparison schools over three years, which was 91.47 percent.

* Statistically significant at $p \leq .05$. Effect sizes are shown in parentheses.

Exhibit 4.6 presents **credit accumulation** at the end of ninth grade for one-year persisters. (These figures include high school credits accumulated in middle school. The numbers of high school credits accumulated by NCHS and comparison students prior to high school were virtually identical.) Credit accumulation was significantly higher among ninth-grade NCHS students in all six group/cohort combinations. Effect sizes were highest for students admitted in 2003-04 (those enrolled in Group 2 schools).

Exhibit 4.7 presents longitudinal credit accumulation patterns. Again, NCHS students performed at higher levels than did comparisons. Specifically, NCHS two-year persisters and three-year persisters outperformed comparison-group students. The differences were statistically significant, and effect sizes were larger for students admitted in 2003-04.

Exhibit 4.6
For Paired Comparisons, One-Year Credit Accumulation

Sample Size (n=1630)

	C1	C5	C9
Group 1 (schools starting in 2002-03)	169	146	138
Group 2 (schools starting in 2003-04)		426	275
Group 3 (schools starting in 2004-05)			476

NCHS Average Credit Accumulation

	C1	C5	C9
Group 1	11.3* (0.32)	11.0* (0.57)	11.0* (0.32)
Group 2		11.6* (0.62)	11.2* (0.23)
Group 3			11.1* (0.31)

Comparison Group Average Credit Accumulation

	C1	C5	C9
Group 1	10.4	8.3	9.5
Group 2		8.9	10.2
Group 3			9.7

Exhibit reads: On average, NCHS ninth-graders in Cohort 1 (n=169) accumulated 11.3 credits. This was significantly more credits than the 10.4 accumulated on average by their matches in comparison schools.

* Statistically significant at $p \leq .05$. Effect sizes are shown in parentheses.

Exhibit 4.7
For Paired Comparisons, Two-Year and Three-Year Credit Accumulation

Sample Size (n=526)

	C1 (3 years)	C5 (2 years)
Group 1 (schools starting in 2002-03)	105	109
Group 2 (schools starting in 2003-04)		312
Group 3 (schools starting in 2004-05)		

NCHS Average Credit Accumulation

	C1 (3 years)	C5 (2 years)
Group 1	33.4* (0.26)	21.5* (0.56)
Group 2		23.4* (0.57)
Group 3		

Comparison Group Average Credit Accumulation

	C1 (3 years)	C5 (2 years)
Group 1	30.8	17.3
Group 2		18.9
Group 3		

Exhibit reads: On average, NCHS students in Cohort 1 (n=105) accumulated 33.4 credits over three years. This was significantly more than the 30.8 credits accumulated on average over that period by their matches in comparison schools.

* Statistically significant at $p \leq .05$. Effect sizes are shown in parentheses.

Exhibit 4.8 compares **promotion** rates, based on (1) the number and percentage of 2002-03 ninth-grade entrants who persisted for three years and were promoted to the eleventh grade or beyond by 2004-05 and (2) the number and percentage of 2003-04 ninth-grade entrants who persisted for two years and were promoted to the tenth grade or beyond by 2004-05. We note again that this statistic is constrained to multi-year persisters, who are a select group of students. Differences between NCHS and comparison students' performance on this standard were considerable. Just over 91 percent of Cohort 1 NCHS students were promoted to eleventh grade or beyond, and about 63 percent of comparison students were promoted to eleventh grade or beyond, a difference of 28 percentage points. Exhibit 4.8 shows that even larger differences were found for Cohort 5 students, who were those admitted in 2003-04. All differences were statistically significant.

Exhibit 4.8
For Paired Comparisons, One-Year and Two-Year Promotion Rates

Sample Size (n=526)

	C1 (2 years)	C5 (1 year)
Group 1 (schools starting in 2002-03)	105	109
Group 2 (schools starting in 2003-04)		312
Group 3 (schools starting in 2004-05)		

NCHS Promotion Rates

	C1 (2 years)	C5 (1 year)
Group 1	96 (91.4%)*	105 (97.2%)*
Group 2		303 (97.1%)*
Group 3		

Comparison Group Promotion Rates

	C1 (2 years)	C5 (1 year)
Group 1	66 (62.9%)	68 (62.4%)
Group 2		214 (68.6%)
Group 3		

Exhibit reads: 91.4 percent of NCHS students who entered the ninth grade in the fall of 2002 and were still enrolled and active in the school in 2004-05 were promoted to the eleventh grade or beyond by the end of the 2004-05 school year. This was a significantly higher promotion rate than that attained by their matches in comparison schools; 62.9 percent of comparison group students were promoted to eleventh grade or beyond.

* Statistically significant at $p \leq .05$, $df=1$. Percentages are shown in parentheses.

Exhibit 4.9 presents longitudinal data on **Regents test performance,** based on mean accumulation of tests with passing scores of 55 or higher by Cohort 1 persisters (students who completed their third year in a sample school in 2004-05), Cohort 5 persisters (students who completed their second year in a sample school in 2004-05), and Cohort 9 persisters (students who completed one year in a sample school in 2004-05). Results on this measure were mixed. For three of six group/cohort combinations,

there were no differences in the performance of NCHS and comparison students. NCHS three-year persisters lagged behind comparison students, and the difference was statistically significant. NCHS students in Group 2 schools (two-year persisters) had higher credit accumulation than comparison students, and the difference was statistically significant. This was true as well for Group 2 one-year persisters.

Exhibit 4.9
For Paired Comparisons, Regents Exams Passed

Sample Size (n=1415)

	C1 (3 years)	C5 (2 years)	C9 (1 year)
Group 1 (schools starting in 2002-03)	105	109	138
Group 2 (schools starting in 2003-04)		312	275
Group 3 (schools starting in 2004-05)			476

NCHS Mean Regents Exams Passed

	C1 (3 years)	C5 (2 years)	C9 (1 year)
Group 1	3.70	1.91	0.33
Group 2		2.59* (0.30)	0.87* (0.43)
Group 3			0.37

Comparison Group Mean Regents Exams Passed

	C1 (3 years)	C5 (2 years)	C9 (1 year)
Group 1	4.59* (-0.34)	1.64	0.25
Group 2		2.04	0.45
Group 3			0.39

Exhibit reads: On average, NCHS students in Cohort 1 (n=105) passed 3.70 Regents exams over three years. Over the same period, their matches passed an average of 4.59 Regents exams, which was significantly more.

* Statistically significant at p≤.05. Effect sizes are shown in parentheses.

Analysis of **suspensions** is based on the 1,632 pairs of students who were enrolled in sample schools in 2004-05 and who entered as ninth-graders between 2002-03 and 2004-05. From this group, in 2004-05, 106 NCHS students were suspended (7 percent) and 75 comparison-group students were suspended (5 percent). Suspension rates were higher among ninth-grade NCHS students than comparison students, and the difference was statistically significant. In 2004-05, suspension rates were also higher among NCHS students than comparison-group students persisting for multiple years.

V. FINDINGS FROM SITE VISIT INTERVIEWS AND OBSERVATIONS

Following passage of the No Child Left Behind Act, several prominent organizations and individuals made a case for extending the term *scientifically-based research* to methods other than those associated with experiments and probability surveys (American Educational Research Association, 2003). Maxwell (2004) elaborated this point of view, arguing that qualitative methods could be effectively used both to verify whether a program caused an effect and to clarify how it caused an effect.

PSA researchers embraced the opportunity to collect qualitative data as part of the NCHS evaluation for the reasons cited by Maxwell—verification and explanation. The verification task led us to ask about the diffusion in NCHS schools of conditions that are empirically associated with better than expected high school outcomes. The explanation task led us to ask about relationships between key NCHS program features and outcomes.

Designing an approach to verification was relatively straight-forward. The literature is replete with research contrasting schools and classrooms that produce high test scores with those that do not. Productive schools are consistently found to be safe and orderly, to emphasize academics, to employ strong leaders, to hold high expectations, and to evaluate student progress frequently (Firestone, 1991; Reynolds, Creemers, Stringfield, Teddlie, & Schaffer, 2002; Teddlie & Reynolds, 2000). Productive classrooms are well-managed classrooms where students concentrate on high-value academic tasks and receive regular feedback (Emmer, Evertson, & Worsham, 2003; Evertson, Emmer, & Worsham, 2003).

Enrollment size is also a factor associated with high school outcomes. Students at small schools participate in extracurricular activities at higher rates than students in large high schools (Cotton, 1996), and students who are more involved in extracurricular activities have better health and academic outcomes (Steinberg & Allen, 2002). The rule of thumb is this: as poverty increases, school size should decrease (Howley & Bickel, 2000).

Following the requirements of the evaluation and leads found in the literature, we framed the qualitative inquiry around three questions: Were NCHS schools small, safe, and focused on instruction and youth development? Did advisories, curricular themes, partnerships, and professional support—matters of particular interest to NCHS sponsors—contribute to effectiveness? What obstacles did NCHS schools face?

Having focused the qualitative inquiry, there were three additional sets of challenges to address. The first set related to data collection and included developing instruments, guaranteeing confidentiality and anonymity, protecting the rights of under-age students, staffing the project with well prepared observers and interviewers,

collecting data in person in a large number of sites, and verifying the quality of the data collected.

The **data collection process** resolved as follows. Samples of NCHS schools were visited in each year of the assessment. By the end of the 2004-05 school year, the 12 NCHS schools that opened in 2002-03 had been visited three times. Eight of the 18 schools that opened in 2003-04 had been visited twice, and the six of 45 schools that opened in 2004-05 had been visited once. Each two-person site visit lasted for at least three person-days, and there were 58 site visits.

PSA interviewers met individually with principals and partner representatives and conducted focus groups with students, teachers, and parents. Teacher focus-group candidates were selected by school-level liaisons, and the groups were typically conducted following school staff meetings. Groups ranged in size from five to eight members and often included all teachers in a school (given the small size of schools). Student focus-group candidates were also selected by school-level liaisons. Active parental consent was then sought for student participation. Only students with active permission were allowed to participate in these well-attended groups, which averaged eight students, and were generally conducted over pizza during students' lunch periods. Parent focus groups were organized by school-level parent liaisons. These groups were smaller than planned, and they typically included the parent association president and two members of the parent association.

During each school visit, PSA observers noted instructional activities for at least 50 minutes in each of three ninth-grade language arts classes. We recorded observations in 10-minute segments taking note of the following: grouping patterns, instructional focus, instructional activities, materials in use, instructional strategy, performance goals, numbers and types of questions teachers posed, and numbers of students on task.

The second set of challenges for the qualitative analysis related to **methods of data analysis**. We were confronted with more than one thousand pages of field notes. To code these data, we formed a team of analysts, all of whom had participated in school site visits. Analysts initially interrogated the data for information relevant to the pre-existing topics specified earlier. Thereafter, the approach to analysis was inductive (Patton, 2002). Analysts met with the qualitative data team leader to identify themes and categories and went back to reexamine the data to verify the presence or absence of those themes and categories. Codes were recorded using NUD*IST, version N6, a product of QSR International.

The third challenge was **presenting the data**. We were interested in communicating in as direct a manner as possible, but we could not identify schools, partner organizations, or individuals by name or by implication, given confidentiality agreements. The qualitative team resolved ultimately to categorize, type, and group NCHS schools rather than to present individual case studies or even partial portraits.

Some of the specificity and power of the qualitative data was sacrificed to this decision, but to a greater good.

Summary

To summarize findings from site visits briefly, NCHS schools in our research sample had many features associated with academic effectiveness. All were small. Almost without exception, they were safe and academically focused. And 75 percent of the schools were rich with instructional supports and youth development activities.

We found considerable diversity among the schools with regard to their thematic integration, teachers' assessments of available professional development activities, and the quality of partnerships with nonprofit external organizations. Finally, we found that advisories, while sometimes helpful, were frequently under-realized either for lack of a curriculum or for lack of staff training.

Students and staff perceived the following environmental conditions as interfering with the optimal development of their schools: enrollment growth based on the addition of successive grades, crowding in shared and limited school facilities, over-the-counter admissions, and uncertainty about the continuance of private funding for programs and initiatives.

Below organized by topic is a digest of the information we gathered in in-depth interviews and focus groups. Informants included students, teachers, principals, and nonprofit partners in 20 schools.

Adoption of Proven Practices

School Size

All NCHS schools that we visited were small (fewer than 400 students each). Almost all parents, students, and teachers viewed the small size of these schools as a great advantage. In the words of one parent, "I was looking for the smallest high school I could find. You know three or four thousand kids at some schools and there is hardly any control." In the words of another parent, "[…the school] had to be small. The teachers here call me. I don't think you get that at a big school. My daughter is doing better at this school [than at any previous school]." Student comments ran along these lines: "Teachers know your name, and it's so small that there is not much trouble here." "Teachers here are like a family; they understand where we are at; they help us with our work; they know we have problems….In other schools, teachers don't care…." "We have a one-on-one relationship with our teachers."

While generally regarded as an advantage, small enrollments also had a downside according to some students. Asked what would happen if he missed a class

or if his grades started to slip, a student noted, "Oh, yeah, my teacher would notice. He notices everything." Another student said, "Teachers get on your nerves because they know everybody."

Teachers and students tended to lament even the inevitable enrollment increases that came with the addition of successive grades. As one teacher said, "When you start out so small, it is almost a family. It started around 90 kids. Now with more kids, there is a closeness that is gone."

The scale of NCHS schools meant that they generally shared buildings, and according to several respondents, this undercut certain advantages associated with smallness. One teacher said, "The number of small schools is expanding. My fear is that replacing [the old school] with small schools with combined enrollments of 3,000 kids…will bring some of the same issues." Another teacher said, "With large groups of kids, you'll always have problems."

Safety

In the overwhelming majority of NCHS schools, students said they felt completely safe. This was true even in neighborhoods where gangs and drugs were prevalent. Typical comments were as follows, "We are a big family. We are like brothers and sisters. We have our arguments, but at the end of the day, we are still friends. It is hard not to get along with anyone at this school."

NCHS principals took special care to insulate students from the turmoil in host schools, especially on Impact School campuses, but there was *one* NCHS school in which intra-building tensions were making safety a concern for students. Students in this school estimated that one or two fist fights broke out every week with students in other schools that shared the same building.

Teachers and principals were worried about intra-building tensions in three additional school buildings. They noted the following: "We have issues…that come up because we have to share the auditorium, library, gym and cafeteria." "There is tension because the boys come down and hit on the girls or try to start fights with the other boys." Lunchroom problems were so widespread that virtually all of the schools sharing campuses had switched to separate lunch periods.

Respondents frequently held building security personnel responsible, at least in part, for problems with shared facilities. The parent coordinator in one school explained, "The school safety officers only see one principal from the host school, and if there are safety issues, the small-school principals do not get the same level of respect."

Reporting on a conversation with a security officer, one teacher ventured that at least some security problems stemmed from limited resources. She said the security

officers in her building said that they couldn't make a separate entrance available to the small schools because they didn't have the person-power to secure multiple entrances.

Instructional Focus

Observations of instruction in 2004-05 were conducted in classrooms that almost certainly represented examples of best practice, since these observations were arranged by school personnel. The data show that almost all students were on task most of the time; specifically, that 90 percent of students were on task during 65 percent of the 620 classroom observation segments conducted by PSA observers. In about 20 percent of observation segments, 60-90 percent of students were on task. In about 15 percent of observation segments, less than 60 percent of students were on task. Instruction took place, for the most part, in whole groups (67 percent of segments). Often teachers gave students information or asked questions (74 percent of segments). In many segments, fact-based or procedural questions were asked (51 percent). In some segments, inferential questions were asked (30 percent) and in some segments, relational questions were asked (21 percent). One or another student was corrected for a disciplinary infraction during many segments (39 percent).

In interviews, teachers reported that they typically viewed instructional focus, rigor, assessment, and support against the backdrop of Regents exams. These comments led us to classify schools instructionally by their approach to Regents. Only four of our 20 sample schools were *not* focused on Regents preparation. One school within the group of those not focused on Regents was designed as a temporary stopping place for students transitioning from a non-academic setting and not as a diploma-granting institution. One of the other three featured learning by doing, and the remaining two schools served students with especially weak basic academic skills. Three of these four schools had a clear instructional focus, although not Regents-related. One appeared unfocused.

In the second and much larger category, we found schools that had accepted, if not yet fully embraced the Regents exams as a way of focusing their curriculum. Nine were fully Regents-focused, two were in transition, and five were struggling to integrate Regents preparation with their content theme.

Among schools focused on the Regents, several had integrated Regents preparation into a well-rounded and engaging curriculum that was not particularly test-centric. An English teacher, for example, noted that the skills she was teaching and the way she was challenging her students would prepare them for their Regents exam even though she did not focus explicitly on the test. In many, if not most, NCHS schools, however, teaching to the Regents was viewed as teaching to standards, not teaching to the test. In other words, teaching to Regents was nothing to be ashamed of or to avoid; it was simply a way to motivate students and to develop in them the skills necessary for success in college. In only one NCHS school did teaching to Regents standards

appeared to be headed in the "drill and kill" direction. This school reminded us that focus is no *guarantee* of excellence; it is possible to focus too narrowly.

Support for Learning

Most NCHS schools offered Regents preparation after school and/or in Saturday academies in addition to the help provided by teachers during regular class time. Many schools focused regular class time explicitly on Regents preparation toward the end of the year. But these observations substantially understate the ways in which NCHS schools worked to meet students' learning needs.

We identified three types of NCHS schools with regard to their level of instructional and personal support: (a) systemically supportive schools (two of 20), (b) generally supportive schools (13 of 20), and (c) inadequately supportive schools (five of 20).

In systematically supportive schools, administrators and teachers were in regular communication with parents and teachers. The schools provided students and their families with access to social services, and this compensated for the fact that advisories were of varying rather than uniformly high quality. Most importantly, in these schools, the general curriculum was tailored to students' needs, and lesson plans and special classes were developed according to students' learning abilities.

In generally supportive schools, despite efforts to provide the assistance described above, circumstances got in the way of responsiveness at times. Some teachers had difficulty individualizing instruction in the classroom. Pointing to this problem, one student observed, "If you raise a hand and say you don't understand something, they say you should have been paying attention. They want you to come after school. Every time…you don't understand something, they say, 'come to tutoring after school.'"

Often a school was categorized as generally rather than systematically supportive because advisory periods were seriously under-utilized. In generally supportive schools, despite some effort to structure advisory periods, the curriculum could be weak. Some teachers were personally unprepared to lead advisories; some teachers failed to follow the advisory curriculum. In schools where advisories were conducted by counselors, some counselors failed to be in systematic contact with teachers about students' progress and needs.

The problems in inadequately supportive schools were at their instructional core. This group included schools in which students and teachers reported that classes were disorganized. It also included schools in which principals or teachers were not viewed by students as skillful despite their willingness to help.

Illustrating student complaints at the far end of the inadequate spectrum were comments such as the following, "I think this place is not organized….When we try to talk to someone about our problems, we are not heard. [The principal] ignores questions when we try and talk to him. I don't know how the principal became a principal because he is not qualified."

What we heard, however, from students and parents in 75 percent of the schools was quite different. As one student said, "Teachers do home visits and call the house for good things and bad things. They call when you are good. You really bond with them…and because you are in such small classes, the teachers really get to know you."

And as another student remarked, "The school provides lots of openings [for students to pursue interests]. They help us and don't make fun; they are patient. The teachers care about our work. They want to make you succeed. They will stay on you until you are doing your work."

In the NCHS schools identified above as systematically supportive, the social-emotional needs of youth and their parents were addressed as well as students' academic needs. In the words of one partner, "If there [are families] that need preventive service, they have access to all the resources of our programs…."

Youth Development Activities

Like individualized educational supports, youth development activities were widespread in NCHS schools. Fifteen of the 20 NCHS schools we visited twice had a range of well-used extracurricular activities that kept students involved and connected with learning and their local communities.

Many youth development activities took place during the school day because they were built into the curriculum. Activities during the school day included year-long and summer internships, field trips, college tours, leadership programs, student government, sports, wilderness retreats, and community service.

After school, there were clubs in which students could participate. Club offerings in one school included guitar, drumming, modern dance, art, and drama. In another school, offerings included drama, swimming, chess, karate, break-dancing, computers, and knitting.

Content Themes, Partnerships, and Professional Support

Stakeholders had a particular interest in the way advisories, content themes, partnerships, and professional support opportunities were playing out in NCHS schools. These reform-focused activities featured prominently in NCHS program theory. The

expectation was that over time themes would render instruction more relevant, that partnerships would provide intellectual capital as well as direct support to staff and to students, and that professional support/development would deepen the practice of teachers and administrators.

Content Themes

We defined a well-executed theme as one that had been seamlessly integrated with a school's mission and curriculum and that provided focus and direction to both. Guided by this definition, we distilled four types of NCHS schools: (a) schools without a theme (five of 20); (b) schools with very limited thematic integration (five of 20); (c) schools that used their content theme outside the classroom in after-school activities, advisory classes, and internships (six of 20); and (d) schools that were successfully integrating their content themes into their academic program (four of 20).

Interestingly, several schools *without* content themes appeared to be as instructionally sophisticated and effective as those with content themes. The schools that seemed most worrisome were those that purported to have themes but had largely failed to develop them. This could, of course, confuse and disappoint students. In one such school, the principal and partner couldn't frankly recall the theme, and they ultimately recollected it differently. Most principals in schools with weak themes had better recall but limited ambition to change the status quo, as illustrated by the following comments: "We haven't worried a lot about the theme. The theme seemed much more important during the planning phase." "The theme has to flow naturally. I don't want to force it. There are pockets [where it is occurring]."

In schools that had applied their themes mainly to after-school programs, internships, and the like, staff members generally experienced a conflict between implementing the theme and following district regulations (especially with regard to Regents preparation, Ramp Up, and teaching English as a second language). This concern was expressed in the following comments: "We must either break from the city's mandates to provide double periods or have an extended school day to accommodate theme-related classes." "It's an ongoing struggle. Every student has to take a double period of math and a double period of literacy. The day is chock-full of requirements…so there's no time to add electives."

While some schools succeeded without themes, schools with integrated themes were among the most vibrant and academically challenging we saw. In these schools the theme "hooked" students. It was their reason for applying to the school and their reason for attending regularly. The theme defined the context for instruction, including Regents preparation.

The following illustrate some of the ways thematic integration was approached. In a drama-focused school, English, history, and theater classes were organized around plays students read and performed, which changed every six weeks. In a school with an

international focus, students took three different languages, and all traveled internationally before graduating. In a school focused on law, Forensic Science replaced the Living Environment as the ninth-grade science class, and teachers adopted mock trials as an instructional strategy.

The mechanics of achieving instructional coherence in schools with effective themes was relatively simple. Teacher communication, professional development, and lesson planning systematically revolved around both the schools' content themes and Regents requirements, as illustrated by the following comment, "Last semester the program director and I worked on a New Visions grant to come up with something to do with literacy and the theme…so I was teaching to the New York State standards of writing reports and informational reading and also incorporating the theme."

We found no relationship between thematic integration and the categories into which NCHS schools could be readily sorted. There were no differences, for example, in the level of thematic integration by the year in which schools opened, the borough or DOE region in which they were located, the nature of their content theme, or the extent to which partners actively supported the theme. Individual processes, personalities, and circumstances appeared to be dominant factors in determining thematic integration.

We did find, however, an unfortunate negative interaction between attention to themes and the new high school admission system. In interviews conducted in 2004-05, principals said they received a higher proportion of students who were uninterested in their schools' themes in that year than previously (even though all admissions processes were designed to take student preferences into account). This problem appeared to have exerted a greater negative impact on schools with strong thematic integration than on other NCHS schools, because the strongly thematic schools could less easily accommodate students with different academic interests.

Partnerships with Local Nonprofits

NCHS schools were developed in partnership with local nonprofits that stepped forward to play this role. The expectation was that partners would bring intellectual capital as well as direct technical assistance to the small-school effort. Partner organizations were also the fiscal agents for the financial contributions of sponsoring grantmakers to NCHS schools.

Partnerships in 2004-05 were deeper than before. About one-third of the schools that PSA researchers visited had very strong partnerships. About one-third had partnerships that were successful in specific areas of school life. And about one-third had ineffective partnerships.

Where there were strong partnerships, the school and nonprofit organization shared a common vision, the partner had a staff member on site in the school, and the partner was regularly involved in most aspects of school life. In the words of a teacher

describing such a partnership, "They're everything. It's like this [clasps hands together]…we wouldn't have a school without them." Describing a strong partnership in another school, a second teacher said, "It's like you have access to a personal adviser. Because he's looking from the outside, he has a great perspective."

Strong partners, while systematically involved in the life of their schools, typically specialized in particular areas of work. One partner provided a part-time coach who served as a sounding board and advisor for the principal. Another partner fostered community involvement and coordinated special programs. Still another partner offered ongoing professional development to teachers.

In the second group of partnerships, community groups played a discrete role or set of roles in the schools rather than a systematic role. For example, they led or co-led advisories, conducted after-school and tutoring programs, involved students in community advocacy, conducted off-site educational programs, and provided social services. One partner provided an artist in residence.

In the third group of partnerships, collaboration was deemed ineffective. This group included arrangements in which community organizations had unclear roles, or their work in the school was inconsistent and needed to be monitored, or where the school and community group differed over the design and implementation of activities, or where personalities were at odds. Such conflicts led one partner organization in our research sample to withdraw its full-time program director.

Again, we found no systematic relationship between the quality and extent of partnerships and the categories into which partners could be readily sorted. Some nonprofits partnered with multiple schools, and even in these circumstances, the quality of partnership activities varied considerably from school to school.

Professional Support

NCHS schools received support from the Regional Office of Small Schools (ROSS), the local instructional superintendent's staff (LIS), the DOE Office of New Schools, and New Visions for Public Schools. Not surprisingly, principals and teachers had difficulty distinguishing among support providers, particularly among providers directly associated with the school system. For this reason, we have organized comments around themes rather than assistance providers, although we name providers when we are certain about the reference.

Trust was a matter of considerable concern in matters related to professional support and development. When staff believed they were receiving help from Department of Education providers, they were eager to understand how providers fit into the hierarchy. In the words of two principals: "At first I wasn't sure what they were supposed to do; how they were going to function; who they reported to. Ultimately, we needed to be able to talk to them about issues we can't talk to our local

superintendent about." "It's important that [he] not report to the superintendent on everything [he] sees and hears."

Principals were pleased with the support they received from mentors. They experienced these individuals as knowledgeable about practical matters and as trustworthy. In the words of one principal, "I have a mentor….I discuss [with him] issues relevant to the mechanics and the politics of running a school. He is an excellent sounding board [for] scheduling concerns and personnel issues."

Having the ear of the DOE hierarchy was a matter of considerable importance to principals when it came to school policy. After acknowledging the help of New Visions, one principal noted, "…but it's the assistance from the person in charge of our schools I need…to talk about issues of control." Referring again to New Visions, another principal said, "When they can, they've been wonderful with professional development, but I don't [go to] them for policy issues."

Networking opportunities seemed to address some of the complexities of school professionals' and partners' needs for access to school officials, problem-solving with peers, and nonjudgmental support. A partner observed, "It helps that New Visions… attempts to bring people together from within the same schools, schools within host schools, and larger communities in the form of partner breakfasts." Networking was seen both as a chance to learn from people in similar situations and as an opportunity to influence the formal school hierarchy. Insofar as school officials attended networking events, respondents believed they would see how problems were clustering.

It appeared from teacher comments that their access to staff development was uneven across NCHS schools. A teacher in one school acknowledged that there were more professional development opportunities in 2004-05 than earlier. She described multiple planning opportunities and retreats that were designed, in her words, "to make ourselves fit the things we say we are." A teacher in another school had not heard about any professional development opportunities available through New Visions. She said, "I'm never alerted about any professional development from New Visions. I don't know if they do it." This teacher relied on a mentor from her student teaching days for help.

Teachers who participated in staff development were generally responsive to opportunities to work for extended periods of time with peers, whether in school or out of school, although staff retreats were a particularly prized. Having attended a weekend retreat, a math teacher remarked, "It was awesome. We mapped out the math curriculum, and we spent the next day mapping out the science curriculum for the next two years in that one weekend."

Obstacles

Interview respondents expressed frustration with the city's new admissions system, with the routines and disciplinary problems accompanying shared space, and with increasing enrollments. Many were also concerned about the sustainability of key program features once grant funds were expended.

Admissions

To promote equity in admissions, at the beginning of the 2002-03 school year, the DOE redesigned the high school application and admissions process to optimize what it described as "informed choice." Under this policy, students are asked to list up to 12 high school choices in priority order, and they are matched to one of those choices. A downside of this redesign from the standpoint of many NCHS school leaders was a reduction in their perceived influence in student selection. According to one teacher, "In the first year, students generally got their first choice of schools. Now, we're getting kids for whom this isn't their first or second choice. So, lots of kids walk in with an attitude." According to one principal, "We are getting kids…on suspension, transfers, and a lot of kids who don't want to be here. These children have a problem with the dress code and are not academically inclined in terms of the focus of school."

Principals seemed to have a common perspective about needed changes. Summarizing the views of many, one said the admissions system should "…just get kids where they want to be. I have a problem with 10 kids in the tenth grade—they just don't want to be here, no interest in [the theme] and they want to get out. Send kids where they want to be; [they should] find ways to make better fits, but that doesn't really happen." School system officials noted, "The high school admissions process is a student-driven process." Students are asked to list *only* schools whose themes interest them, but it appears that sometimes they do otherwise.

A related problem concerned "over-the-counter" students. The DOE defines "over-the-counter" students as those who do not participate in the high school admissions process (or who request a transfer for medical or safety reasons or due to travel hardship) and are assigned to schools with available seats based on geographic proximity and interest. According to NCHS respondents, a high proportion of students admitted over the counter were not, in fact, interested in NCHS themes or in the special practices sometimes found in these schools, such as wearing uniforms.

In addition to causing disruption, these students were believed to undermine the schools' attendance records. A teacher explained, "We have had a lot of long-term absences this year. The over-the-counter [student] has had a big impact. They are trying to find seats for students who have no interest in our school…which results in their absences. You cannot take a student off your register unless you know they have been enrolled in another school. Sometimes parents leave the country, and they stay on your register."

Sharing Facilities

Of the 17 schools in our research sample located on a larger high school campus or shared facility, only three had strong positive campus-wide relationships. Six reported mixed relationships—some good, some bad—whether with other small schools on their campus or with the host school, and eight had little or no contact with other schools on their campus. At the time of our site visits, the school system was introducing "building councils" in shared facilities to increase collaboration, but those councils had not yet achieved much traction, and there were problems.

As noted earlier, sharing facilities presented challenges for school security. Schools in the same building often had different discipline codes and expectations for student conduct. It also made scheduling to use theaters, gyms, and lunchrooms in the words of one teacher "a programming nightmare." In a few small schools, staff reported that they did not even have reliable access to their school buildings. A principal said, "I don't have a key to this building, and I don't have a parking space…." A teacher in another school explained, "Sometimes we get trapped outside the building, and we can't get in."

Students complained that housing schools with different schedules in the same building led to distractions in class. There was noise associated with the beginning and end of separate school days and with the changing of separately scheduled classes. Added to this was the intentional disruption caused by some students. One student remarked, "I'm tired of other schools coming up here. [They] overflow the toilets with towels. They knock on [classroom] doors and run."

Some schools sharing facilities had shared schedules, and this too could cause problems such as logjams at the school entrance in the morning. As one student said, "There are too many kids coming into one entrance. If all the buses get here at the same time, it takes a long time to scan your ID and get upstairs."

Summing up the problem with sharing space, a student simply noted, "We need our own building and our own space. Sometimes you just want your own room." In one way or another, in every school we visited that occupied a shared facility, this sentiment was expressed. At times it was accompanied by an acknowledgement that reconfiguring high school facilities to accommodate small schools would come at a cost. Noting this, one principal said that his colleagues should make more of a fuss about their buildings. In his words, "Realizing complexities didn't keep me from pushing for what my school needed."

Increasing Enrollments and Crowding

According to some principals, over-crowding exacerbated the problem of sharing facilities. One principal observed, "…we only have seats for 108. We were

told we had to offer seats to everyone. I am supposed to have 108, and so far 180 kids have been accepted."

Principals were not alone in this complaint. Some teachers reported having an unexpectedly large number of students and not enough classroom space. One teacher noted, "The fact that I need to teach outside of the school in the trailer alludes to the…issue…."

Teachers were quick to point out the relationship between crowding and instruction. In the words of one, "We are supposed to take in more students next year, but we do not know how to handle it. When you go up by eight to 10 kids per class it just affects the whole model of the school. What we are going to get is an overcrowded small school."

Parents too were concerned about crowding and class size. In the words of one parent, "They said they were working with small classes, but [when] 20 and 25 students [are in a room it] is not a small class." A principal noted that classes averaged 28 students when 20 would be more appropriate.

Funding

Through their lead partner, each New Century High School received $400,000 over a four-year period for expenses associated with start-up. Several principals and partners echoed this principal's sentiment regarding the impending loss of funds: "Money is going to be a profound issue once the grant [from New Visions] finishes up."

Some respondents focused on the sustainability of partnerships. A partner said, for example, "We can provide other things but without money it will really be a loss."

Other respondents focused on the sustainability of instructional programs. A principal said, "[Our] school had to change pedagogy based on the fact that… money…will run out."

Several schools and partners had begun to look for replacement funds with support from New Visions, but for other partners, the dilemma was as follows: "The start-up money is mostly gone. The sustainability plan is not there. We can maintain our effort for the coming school year but beyond that we cannot project our commitment."

VI. CONCLUSIONS

We hope this report promotes active discussion among NCHS stakeholders and reformers with diverse visions for improving schools. The data strongly suggest that students benefited significantly from enrollment in NCHS schools, but as with most program evaluation research, the data do not prove unequivocally that this is the case.

Apparently comparable students were more successful in NCHS high schools than in other New York City high schools, but an attribution problem arises in trying to claim that the schools *caused* the outcomes. There are plausible alternative interpretations of the data. Unobserved differences between NCHS and comparison-group students, such as the students' personal motivation and the involvement of their parents, may have influenced their high school achievement.

Below we present the basic argument for a program effect, some counter-arguments, and our attempt to reconcile these points of view. Next, we make recommendations for research and action.

The Case for an Important Program Effect

Available data show that students educated in NCHS schools in 2004-05 were better prepared for graduation than comparable students in traditional schools. All precursors—attendance rates, credit accumulation, promotion rates, and the number of Regents exams passed—pointed in the right direction.

The nature of the contribution of NCHS schools to improved student outcomes appears to have been swift and important (as well as statistically significant). The ninth-grade holdover rate in NCHS schools in 2004-05 was 16 percent. The ninth-grade holdover rate citywide was 28 percent.[6] The ninth-grade holdover rate in schools selected as NCHS comparison sites based on the similarity of their demographics was a whopping 47 percent. The meaning of early success has been demonstrated by research conducted in the public schools of Philadelphia, Baltimore, and Chicago. Ninth-grade holdovers in those cities were five times more likely to drop out than promoted students (Legters, Balfanz, Jordan & McPartland, 2002; Neild & Balfanz, 2001; Roderick, 1993; Roderick & Camburn, 1996, 1999; Simmons & Blyth, 1987).

[6] Citywide ninth-grade promotion rate available at: http://www.nycenet.edu/Administration/mediarelations/PressReleases/2005-2006/02012006pressrelease.htm .

Limits of Analysis

Those who are skeptical that NCHS schools *caused* the observed differences in student performance can point to other possible reasons for outcomes. First, we used a quasi-experimental research design to assess student achievement results. While this design was optimum under the circumstances, it set limits on the clarity of research findings. As noted, it's arguable that student and teacher selection effects were responsible for some (if not all) differences in student performance.

Second, NCHS schools received substantial economic and other supports from private foundations in the years covered by this analysis. Each new school received $400,000 on top of its public funding over a four-year period. It is arguable that the extra resources provided to these schools and not their design features promoted success. How NCHS schools perform when the playing field is level remains to be seen.

Finally, the sample selection process provides some reason for caution, although this reservation is less persuasive than the prior two. Propensity scoring requires complete demographic data for matching treatment with comparison-group students, and students with complete data typically have higher achievement than those with missing data. This was the case with our sample. Our research sample is not a substitute for the complete NCHS population. Herein, we described and compared the performance of a group of better than average NCHS students and their matches in other schools. One could argue that lower achievers might not perform as effectively in NCHS schools as sampled students, but the preponderance of published research suggests otherwise. Typically students with lower initial achievement benefit more from placement in smaller schools (Lee & Smith, 1997), which is to say that our sampling process placed NCHS schools at something of a disadvantage and not the other way around.

Final Arguments

As useful as quantitative data are for estimating possible program effects, it's true, as the first counter-argument notes, that quantitative data often take inadequate account of the complexity of social reality. Unmeasured factors can continue to be at play in any comparison-group study. In this particular case, we conclude, however, that NCHS schools as they existed at the time of this study (which is to say with their additional financial resources) contributed importantly to students' educational outcomes.

The strongest supporting evidence is that achievement outcomes were anticipated by and consistent with implementation data. Achievement outcomes were weakest in 2002-03, when implementation of academic program components was weakest. Achievement outcomes peaked for entrants to NCHS schools in 2003-04, at

which time the measured quality of the schools' instructional climates also peaked. Outcomes dipped in 2004-05 along with implementation.

Further, and importantly, observations and interviews with parents, students, and teachers confirmed what survey and outcome data suggested. NCHS schools were typically instructionally focused. And they applied rigorous standards derived from Regents requirements in framing instructional goals. The schools were also academically and personally supportive. There were differences among NCHS schools with regard to thematic integration, the quality of partnerships, and the implementation of advisory periods, but the basic elements in successful academic programming were generally acknowledged and pursued.

Challenges

NCHS schools were not, of course, picture-perfect. Suspension rates in NCHS schools grew from 2 percent in 2002-03 to 4 percent in 2003-04 to about 6 percent in 2004-05. Suspension rates were highest in NCHS schools on Impact campuses, but the rates in NCHS schools elsewhere were not trivial. The mean NCHS suspension rate in 2004-05 (6 percent) equaled the citywide average. This likely signals a problem. Generally speaking when suspension rates go up, the overall level of students' school connectedness goes down (McNeely, Nonnemaker, & Blum, 2002).

Respondents identified a number of external factors that could have driven disruptive behavior by students in NCHS schools but no internal factors, perhaps because we did not ask directly about factors under their control. Increasing enrollment, whether solely associated with the addition of successive grade levels or complicated by actual over-crowding, was considered a likely cause. According to students and teachers, as enrollments increased, students were less well known; they were less trusted and perhaps less trustworthy; and behavioral problems ensued. Teachers and students unequivocally preferred the quality of relationships in schools with only a couple of hundred students.

Another source of conflict in NCHS schools, according to principals and teachers, was the quality of and control over school facilities. Most NCHS schools followed the "schools-within-a-school" model, meaning that they operated as separate and autonomous units within larger school buildings. As a result, principals needed to negotiate the use of common space and to defer to host schools on matters of safety and building operations. The competition for control and resources could stir resentment among all parties, especially students, and student resentments sometimes boiled over into fights.

A third problem was over-the-counter admissions. This is a practice, according to school system officials, of assigning students who did not participate in the high school admissions process (or who requested a transfer due to documented medical or safety reasons or due to travel hardship) to schools with available seats, based on

geographic proximity and student interest. According to interview respondents, a high proportion of students admitted over the counter were not, in fact, interested in NCHS themes or in the special routines sometimes followed in these schools, such as the requirement to wear uniforms.

Recommendations for Research and Action

We turn briefly now to the problem of assessment. Going forward, it will be important for NCHS sponsors to develop valid, reliable, and cost-conscious ways of evaluating their schools. At a minimum, one would want to avoid false attributions of success and failure. Our experience with this evaluation may serve as a cautionary tale. We found, given the small size of school-level samples, that data needed to be aggregated across schools (i.e. considered at the level of school groups) for us to validate academic outcomes (attendance, credit accumulation, etc.) from comparison-group studies against implementation findings derived from surveys and observations.

This leads us to recommend that NCHS sponsors consider the option of combining achievement data across years to increase the numbers of data points (or observations) when attempting to understand the value added by NCHS schools. We also suggest the use of data beyond student performance measures and, in particular, that assessment processes include teacher surveys. We found that teachers were more effective raters of school conditions associated with achievement than principals, partners, or students.

Our second set of recommendations concerns school size. NCHS students' and professionals' views were aligned with research on the subject of school size, which indicates that social bonds peak in schools with fewer than 300 students. Academic achievement peaks, however, in schools where enrollments range from 600 to 1,200 students (Lee & Smith, 1997), and so, like everything else, one school size doesn't fit all. We suspect that NCHS schools will want to maintain some flexibility regarding projected enrollments.

Our next set of recommendations concerns the practice of over-the-counter admissions. It is one thing if students admitted over the counter to NCHS schools are poorly matched to the schools simply because the students receive incomplete or inaccurate information. Such a problem can be corrected through an enhanced counseling and referral process. It is another thing if high school seats are so limited citywide that appropriate placements cannot be found for average students. While not easy to correct given resource constraints, this problem is easy to frame. It is a third thing if over-the-counter students are students that nobody wants. If that is the case, the school system has a very particular capacity problem—an underserved population whose needs should be better defined and addressed.

Our final concern has to do with any unintended negative consequences levied on the broader school system by the NCHS initiative. It's almost inevitable that the

rapid emergence of new schools would increase competition across the system for resources—good students, professional talent, and building space—and risk redistributing rather than adding to system-wide capacity. How can NCHS sponsors increase the likelihood of the latter?

We have a few ideas. First, future evaluations of NCHS schools should formally consider possible side-effects of the initiative, such as over-crowding. Determining if crowding has increased seems straightforward enough, but there are complexities. The perception of over-crowding appears to be very strong in both the broader school system and in the NCHS subsystem. It is possible that the formula defining over-utilization (as well as security needs) may need to be rethought in the context of shared facilities.

Second, given increases in the number of small schools across the city, evaluations of small high schools should take outcomes for the entire population of New York City high school youth into account. We have been guilty of pitting NCHS schools against traditional high schools by comparing the two to suggest the superiority of the smaller schools. In the long run what matters, of course, is that drop-out rates decrease and educational achievement increase throughout the system.

Third and finally, NCHS sponsors should continue to focus on improving pedagogy and thereby learning. Even if the apparent success of NCHS schools was considerably shaped by student and teacher selection effects, which we doubt, it remains true that teachers experienced these schools as facilitating. If teachers' readiness to grow is met by powerful professional development opportunities, city schools will surely benefit as NCHS teachers migrate over time throughout the system.

There are two well-known explanations for the difficulty that urban students face in their first year of high school. One posits that students with weak academic skills are too often placed with teachers who are unprepared to help them master the basics (Balfanz, McPartland, & Shaw, 2002; Neild & Balfanz, 2001; Roderick & Camburn, 1999). The other posits that students get knocked off balance by the stresses and opportunities of the larger and more complex social environments found in high schools (Lee & Smith, 2001; Roderick & Camburn, 1999). We conclude on the happy observation that in NCHS schools we may have a way of addressing both of these problems.

References

American Educational Research Association. (2003, January 26). *Resolution on the essential elements of scientifically based research.*

Balfanz, R., McPartland, J., & Shaw, A. (2002). *Reconceptualizing extra help for high school students in a high standards era.* Paper presented at the Preparing for America's Future High School Symposium, Office of Adult and Vocational Education, U.S. Department of Education, Washington, D.C.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences.* New York, NY: Academic Press.

Cotton, K. (1996). *School size, school climate, and student performance*. Portland, OR: Northwest Regional educational Laboratory; www.nwrel.org.

Emmer, E.T., Evertson, C.M., & Worsham, M.E. (2003). *Classroom management for secondary teachers*. (6th ed.) Boston: Allyn & Bacon.

Evertson, C.M., Emmer, E.T., & Worsham, M.E. (2003). *Classroom management for elementary teachers*. (6th ed.) Boston: Allyn & Bacon.

Firestone, W.A. (1991). Educators, researchers, and the effective schools movement. In J. R. Bliss, W. A. Firestone, & C. E. Richards (Eds.), *Rethinking effective schools research and practice* (pp. 12-27). Englewood Cliffs, NJ: Prentice-Hall.

Howley, C.B., & Bickel, R. (2000). *Results of four-state study: Smaller schools reduce harmful impact of poverty on student achievement.* Washington, DC: The Rural School and Community Trust; www.ruraledu.org/nat_sum.html.

Lee, V., & Smith, J. (1997). High school size: Which works best, and for whom? *Educational Evaluation and Policy Analysis, 19*(3), 205-228.

Lee, V., & Smith, J. (2001). *Restructuring high schools for equity and excellence: What works.* New York: Teachers College Press.

Legters, N., Balfanz, R., Jordan, W., & McPartland, J. (2002). *Comprehensive reform for urban high schools: A talent development approach.* New York: Teachers College Press.

Maxwell, J. (2004). Causal explanation in qualitative research, and scientific inquiry in education. *Educational Researcher, 33*(2), 3-11

McNeely, C., Nonnemaker, J., & Blum, R. (2002). Promoting school connectedness: Evidence from the National Longitudinal Study of Adolescent Health. *Journal of School Health,* 72, 138-146.

Neild, R., & Balfanz, R. (2001). *An extreme degree of difficulty: The demographics of the ninth grade in non-selective high schools in Philadelphia*. Paper presented at the Annual Meeting of the American Sociological Association, Anaheim, CA.

Patton, M. (2002). *Qualitative research and evaluation methods*. Thousand Oaks: Sage Publications.

Reynolds, D., Creemers, B., Stringfield, S., Teddlie, C., & Schaffer, G. (2002). *World class schools: International perspectives on school effectiveness.* New York: Falmer Press.

Roderick, M. (1993). *The path to dropping out: Evidence for intervention.* Westport, CT: Auburn House.

Roderick, M., & Camburn, E. (1996). Academic difficulty during the high school transition. In Sebring, P. (Ed.), *Charting reform in Chicago: The students speak.* Chicago: Consortium on Chicago School Research.

Roderick, M., & Camburn, E. (1999). Risk and recovery from course failure in the early years of high school. *American Educational Research Journal* 36: 303-343.

Rossi, P., Freeman, H., & Lipsey, M. (1999). *Evaluation, a systematic approach, Sixth Edition.* Thousand Oaks: Sage Publications.

Schiller, K. (1999). Effects of feeder patterns on students transition to high school. *Sociology of Education* 72: 216-233.

Simmons, R., & Blyth, B. (1987). *Moving into adolescence: The impact of pubertal change and school context.* New York: A. de Gruyter.

Steinberg, A., & Allan, L. (2002). *From large to small: Strategies for personalizing the high school.* Boston: Jobs for the Future.

Teddlie, C., & Reynolds, D. (2000). *The international handbook of school effectiveness research.* New York: Falmer Press.

APPENDIX
External Validity of Year 3 Quantitative Analyses

This appendix addresses questions about the external validity of the outcome analyses presented in this report. Specifically, did propensity score matching yield a research group with comparable characteristics to the population of ninth-grade students entering NCHS schools in the three years extending from 2002-03 through 2004-05? We begin by describing the structure of the student-level database and the decision to limit analyses to ninth-grade entrants.

Database Structure

Three confounders needed to be isolated in NCHS outcome analyses: (a) the number of years NCHS schools were in operation, (b) the calendar year that students entered NCHS schools, and (c) the grade assignment of students at entry to an NCHS school. The first two factors could be controlled for during data analyses. Effects relating to students' grade at entry were, however, best controlled for by limiting analyses to ninth-grade entrants. This is because more than 80 percent of NCHS students enrolled as ninth-graders. Exhibit A1 displays the cohort structure of the NCHS database with a cohort defined by the year and grade assignment of pupils upon enrollment in an NCHS high school.

Exhibit A1
Enrollment by Cohort

Grade	Enrollment by Cohort 2002-03	2003-04	2004-05
9th	Cohort 1 N = 900	Cohort 5 N = 2,767	Cohort 9 N = 6,599
10th	Cohort 2 N = 152	Cohort 6 N = 425	Cohort 10 N = 1,401
11th	Cohort 3 N = 33	Cohort 7 N = 39	Cohort 11 N = 65
12th	Cohort 4 N = 15	Cohort 8 N = 5	Cohort 12 N = 102
Special Education	Un-graded N = 5	Un-graded N = 14	Un-graded N = 112

Exhibit reads: Nine hundred students enrolled in NCHS schools in the ninth-grade in 2002-03. These students are members of Cohort 1.

Because the aim of this research was to provide a balanced portrait of NCHS students *across* schools, analyses that captured data for the preponderance of students and schools were substantially preferable to those that represented idiosyncratic cases. We decided, therefore, to focus on the very large sub-sample of students in Cohorts 1, 5, and 9.

Sample

In all, 10,266 students enrolled as first-time ninth-graders in NCHS schools in the three years extending from 2002-03 through 2004-05. The school records of a substantial number of these students (N = 2,731) lacked at least one data element required for propensity matching, and this left us with 7,535 eligible study participants. An appropriate match[7] was found for 1,875 of these students,[8] and all were included in holding power analyses.

Follow-up data were sometimes missing for one or both students in an effectively matched pair. This was typically due to a student's non persistence in a sample school. Performance data for an entire pair was, of course, lost when either a treatment or comparison student left a sample school early.[9] The actual number of participants in most analyses was about 1,630.

Given our interest in comparing the performance of students persistently enrolled in sample schools, a criterion needed to be set for continued enrollment. We defined a student as continually enrolled when the sum of that student's "total days absent" and "total days present" in a sample school was 150 days or more. (School years average about 180 total days).

The 150-day standard meant that we included students with zero days of attendance as long as they remained officially enrolled in a sample school. Our objective in setting this standard was to include the full mix of high and low performers

---

[7] Matches were considered appropriate under the following circumstances: (a) students were identical with regard to race, gender, recent immigrant status, ELL eligibility, free- or reduced-price lunch eligibility, and special education status, (b) attendance differences did not exceed 0.15 standard deviations, ELA scale score differences did not exceed 0.10 standard deviations, and math scale score differences did not exceed 0.20 standard deviations, and (c) age differences did not exceed 180 days.

[8] The odds of finding an appropriate match grew substantially worse over time, for two reasons. First, comparison schools increased their rate of holding ninth-grade students over for a second year in ninth grade, from 39.0 percent in 2001-02 to 42.2 percent in 2002-03 to 45.3 percent in 2003-04. As holdovers grew in numbers, new admissions declined. Second, NCHS schools increased in number and enrollment. The ratio of NCHS to comparison-group students was about 1:10 in 2002-03. In 2003-04, it was about 1:3. By 2004-05 there were only slightly more comparison-group students than NCHS students, and the ratio was about 1:1.

[9] Continued enrollment by a pair was necessary because difference scores formed by subtracting a comparison-group student's score from the score of the matched NCHS student were the unit of analysis.

in analyses but not to include students who had either unofficially dropped out or were enrolled in another school for a substantial part of the school year.

It was critically important, of course, to identify the population of early leavers in sample schools. Holding power is an important indicator of school success. Exhibit A2 presents the progression of data loss, parsing separately cases that were lost due to students' early withdrawal from NCHS and comparison schools. More NCHS students persisted than comparison group students.

Exhibit A2
Sample Attrition

Reasons	N
NCHS students (Cohorts 1, 5, and 9) with data elements required for propensity matching	7,535
NCHS students for whom a matched case was found	1,875
Matched NCHS students who persisted	1,757
Matched comparison group students who persisted	1,731
Total matched persisters	1,632

Exhibit reads: There were 7,535 NCHS students with data elements required for propensity matching.

Given the proportion of eligible students included in the research sample, external validity was a key concern. Exhibit A3 presents data comparing the 1,875 students included in the research sample (a matched case was found) and the 7,755 NCHS students enrolled as first-time ninth-graders and not included in the research sample. Chi-square analyses examine associations between membership in the groups and categorical variables; *t*-test procedures examine associations between membership in the groups and continuous variables.

As expected, students who were effectively matched were less at-risk than NCHS students lacking data for the matching process. Matched students had higher attendance and standardized test scores, and they were younger and therefore less likely to have been held over. These differences were expected because the absence of data is generally associated with lower attendance rates and performance. Given this finding, readers must be cautious in generalizing from the paired-comparison research sample to the population. The academic performance of the population may be different from that of the research sample.

Exhibit A3
Differences in the Eighth Grade between NCHS Research Sample and Other NCHS Ninth-Grade Entrants

Characteristic	NCHS Research Sample			Other NCHS Students			Significance
	N	*Percent*		*N*	*Percent*		^{2}
Female	1875	56.4%		7755	51.7%		13.335**
Race							
Asian	1875	0.7%		8391	4.3%		57.090**
Black	1875	52.3%		8391	38.0%		129.749**
Hispanic	1875	46.4%		8391	45.7%		0.299
White	1875	0.6%		8391	3.7%		48.411**
Other	1875	0.0%		8391	0.7%		12.582**
Recent Immigrant Status	1875	0.4%		8391	8.9%		162.356**
Special Education Status	1875	1.2%		8391	10.4%		161.261**
ELL Eligibility	1875	1.2%		8391	11.6%		189.143**
FRPL Eligibility							
Yes	1875	90.7%		8391	64.6%		488.765**
No	1875	1.3%		8391	5.6%		60.159**
Missing	1875	8.0%		8391	29.8%		380.111**
	N	*M*	*SD*	*N*	*M*	*SD*	*t*
Attendance	1875	0.931	0.045	6797	0.896	0.104	-21.510**
Math Scale Score	1875	705.540	21.125	6979	695.691	39.636	-14.472**
English Scale Score	1875	686.592	15.829	6424	684.536	24.867	-4.289**
Age (in days)	1875	4979.162	148.902	7755	5098.424	278.255	25.538**

*$p \leq .05$, **$p \leq .000$.

Exhibit reads: 56.4 percent of students in the NCHS research sample were female, a rate that was statistically higher than the percentage of females in the group not selected for the research sample.