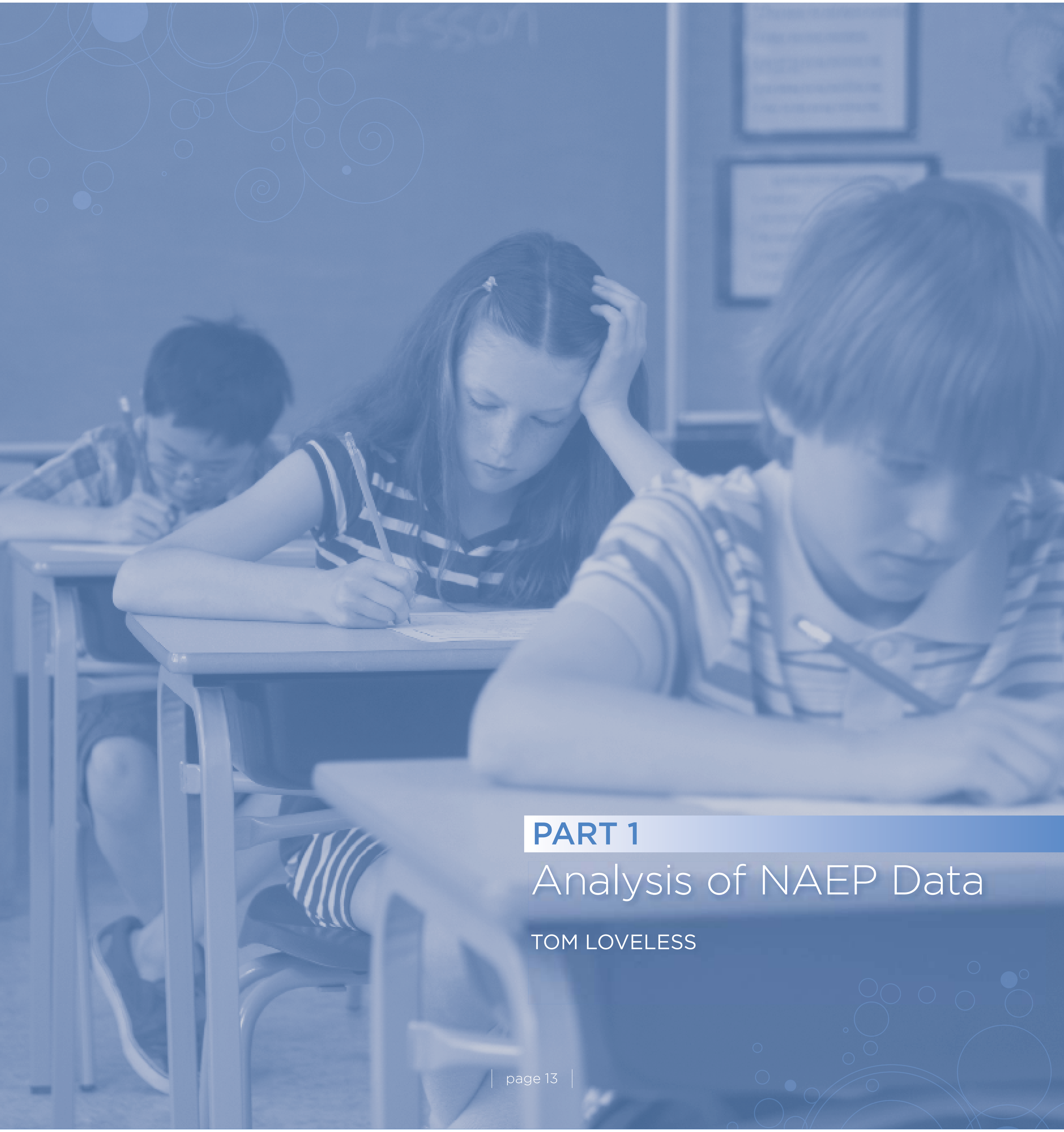


PART 1
Analysis of NAEP Data

TOM LOVELESS

HIGH-ACHIEVING STUDENTS IN THE ERA OF NCLB

In 1972, Commissioner of Education Sidney P. Marland Jr. presented a report to Congress on the education of gifted and high-achieving children in the United States. The Marland Report argued that America had too few challenging programs to meet the needs of its high-achieving students. Just fifteen years earlier, the Russian launch of Sputnik had led to a flurry of programs promoting mathematics and science. Within a few years, however, these programs were eclipsed by a focus on societal inequities—especially those related to race and poverty—and efforts were launched to eradicate similar inequalities in U.S. schools. Gifted programs came under fire for being elitist. Some dwindled away from lack of funding.

In addition to urging that gifted programs address a broad array of talents and abilities, the Marland Report warned Congress that bright minority students are particularly vulnerable:

Intellectual and creative talent cannot survive educational neglect and apathy. This loss is particularly evident in the minority groups who have in both social and educational environments every configuration calculated to stifle potential talent.[1]

Attitudes toward bright children have waxed and waned over the decades. The No Child Left Behind (NCLB) Act of 2001 sought to fuse equity and excellence into a single initiative, promoting academic achievement in the pursuit of equity.[2] Historically, the federal government provided additional revenue to schools serving disadvantaged children, ostensibly so that schools could offer services that would help poor children learn. The architects of NCLB sought to transform the federal education dollar from a school entitlement into an incentive to prod schools towards better performance.[3] Universal proficiency became the nation's foremost education goal.

Incentives shape behavior. Some analysts today express the concern that, by focusing attention on the education of students at the bottom of the achievement distribution, NCLB is surely encouraging schools to neglect high achievers. After all, schools face consequences for failing to move low-achieving students to proficiency. Students in schools that fail to make adequate progress for two consecutive years must be offered the option of transferring to another public school. A school that continues to fall short faces possible replacement of its teaching staff, conversion to a charter school, or state takeover. Nothing, however, happens when schools fail to boost the learning of already-proficient students to higher levels. As Susan Goodkin argued in the *Washington Post*, "By forcing schools to focus their time and funding almost entirely on bringing low achieving students up to proficiency, NCLB sacrifices the education of the gifted students who will become our future biomedical researchers, computer engineers, and other scientific leaders."[4]

Are these concerns well founded? Do the incentives of NCLB create a Robin Hood effect, yielding gains for low-achieving students but at the expense of high achievers? That's what we set out to investigate.

LITERATURE REVIEW

Faced with a powerful incentive to boost the test scores of students on the borderline of proficiency—"bubble kids," as they are sometimes termed—schools might be expected to focus resources on that point in the achievement distribution and neglect the extreme upper and/or lower ends. If such educational "triage" is actually practiced, high-achieving students would lose out by making less academic progress than that of which they are capable. Very little research has been conducted on this topic, but three studies stand out for their sound research methods.

Derek Neal and Diane Schanzenbach examined test scores in the city of Chicago in 2001 and 2002[5], a period when, due to NCLB's impending implementation, the city's assessment regimen shifted from low- to high- stakes testing. They found that students in the middle of the achievement distribution—in particular, those clustered around the threshold of proficiency—made the greatest gains in reading and math. The evidence was mixed as to whether high achievers made the gains that would have been expected based on previous test scores, but the bottom two achievement deciles definitely lagged behind. The same pattern was found in a second batch of test scores from 1996, right after the Chicago school system instituted its own local accountability system. Evidence of educational triage is indicated, but not necessarily at the expense of high achievers. The students losing out seem to be those who are so far below the cutoff for proficiency that they stand little chance of getting over the proficiency bar.

Matthew Springer conducted a similar analysis using data from an entire state.[6] He analyzed test scores from the Northwest Evaluation Association, a national organization that offers assessment services, and focused on the accountability system of a single western state (left unnamed). Springer found no evidence of triage there. Examining test score changes over a three-year period, Springer detected gains across the distribution of achievement. Unlike Neal and Schanzenbach, Springer detected the largest gains among the lowest achievers. But high achievers gained, too. Interestingly, they made gains in schools facing NCLB sanctions—and did not show gains in schools immune from sanctions because the schools had previously made adequate yearly progress (AYP)—the opposite of what one would expect if schools were redirecting resources away from high achievers in response to NCLB's incentives.

Randall Reback examined Texas data from the 1990s in search of signs of triage.[7] He compared the gains made by students in schools facing sanctions under the Texas accountability system with gains made by typical students at similar points in the distribution of achievement. The Texas accountability system at the time based school sanctions on pass rates, much like NCLB. Reback found significant gains by students whose improvement most influenced state ratings, but the scores of very low-achieving students also improved. High achievers did not fare well, and Reback concluded that "relatively high achieving students perform worse than usual if their own performance is irrelevant to the short-run accountability incentives."[8] A cautionary note: the undemanding content of the Texas Assessment of Academic Skills (TAAS), which produced the data analyzed in the study, has been well documented. Some experts describe the TAAS as covering skills and knowledge several years below grade level, which raises questions as to whether it is an adequate instrument for measuring the gains of high achievers.[9]

These three studies yield no clear conclusion as to whether NCLB-style accountability encourages educational triage. In particular, it is unclear how high achievers fare under such systems. They gained (Springer), lost (Reback), and experienced mixed results (Neal and Schanzenbach). In addition to these mixed and inconclusive findings, one of the difficulties in generalizing from studies that focus on a single locale (city or state) is that outcomes may be influenced by other atypical factors. In Chicago, for example, the city's accountability system overlapped with that of Illinois and included a heavy dose of student accountability in the form of mandatory summer school for failing students. Few local accountability systems include strong student accountability, and NCLB is silent on the matter. Because NCLB is a national policy with national implications, an examination of trends in national achievement is informative for understanding how the law may affect high-achieving students.

THE PROBLEM ADDRESSED IN THIS STUDY

The incentives of NCLB are geared towards improving the education of low-achieving students to close achievement gaps. Have low achievers gained the most in the NCLB era? What about high achievers? Data from the National Assessment of Educational Progress (NAEP) are analyzed to compare national achievement trends of low and high achievers. This analysis cannot test causal theories relating to NCLB (or anything else) since NAEP data are cross-sectional, offering a snapshot of how students are performing at a single point in time. However, because NAEP is the only test given to a nationally representative sample each time it is administered, its data give the best estimate of trends in national achievement.

NAEP regularly assesses students in reading and mathematics at fourth and eighth grades. The 10th and 90th percentiles on the NAEP scale are used in this analysis to identify "low achievers" and "high achievers." National averages on NAEP have been going up since 2000. In an environment of rising average scores, what is happening at both ends of the distribution? If the distribution of achievement is shifting upward across all performance levels—all ships rising—everyone would be getting better at about the same rate with no one group having an apparent advantage over another. A compressed distribution or narrowing of the gap between the 10th and 90th percentiles would occur if low achievers gained more than high achievers, or if high achievers' scores declined while low achievers' scores rose. In either case, the bottom would be catching up with the top. A widening distribution, on the other hand, would result if scores of high achievers rose more than those of low achievers, or if low achievers' scores declined while high achievers' scores went up. Regardless, the gap between the two groups would grow larger. High achievers would be outdistancing their peers by even more.

Based on the thrust of NCLB, a plausible hypothesis to begin with is that the distribution of NAEP scores is compressing, with low achievers making gains, high achievers staying flat or even declining, and the achievement gap between the two groups narrowing. After all, NCLB gives schools and policymakers no incentive to boost the scores of high-achieving students. The studies reviewed above offer three reasonable hypotheses about the test scores of high-achieving students: that they went up (Springer), went down (Reback), or were mixed or neutral (Neal and Schanzenbach). One benefit of NAEP is that scale scores run from 0-500, and even the top 10% of scores are immune from a ceiling effect.

RESEARCH QUESTIONS

The study addresses four questions:

1. What has happened to the national NAEP scores of high and low achievers since the advent of NCLB? Reading and math scores at the 10th and 90th percentiles are analyzed for fourth and eighth grades.
2. Was a trend in place before NCLB? National NAEP data prior to NCLB are examined.
3. Is it NCLB accountability or accountability in general that is associated with changes in the achievement gap? State NAEP data from the 1990s are analyzed to compare the gains of low and high achievers in states with and without accountability mechanisms in place before NCLB was enacted.
4. Who are America's high achievers? Student level data from the 2005 NAEP restricted-use files are summarized to paint a portrait of America's high-achieving students. A subgroup of students is singled out for special attention: high achievers who are black, Hispanic, or poor—special subgroups under NCLB.

DATA TREATMENT

Three different NAEP sets of data are used in the analysis—national, state, and student-level restricted-use files. The data addressing research questions #1 and #2 are national means at the 10th and 90th percentiles for students attending public schools. The data in question 3 are 10th and 90th percentile means of state NAEP scores. Question 4 uses student-level data from the restricted-use 2005 NAEP files.

Why use the 10th and 90th percentiles of NAEP to define low and high achievers? An argument could be made to use NAEP's own achievement levels. After all, NCLB sanctions are tied to "proficiency," not to percentiles. Yet the validity of NAEP achievement levels has been questioned since their inception.[10] Moreover, too few students score at NAEP's advanced level—less than 5% in fourth-grade math in 2000, for instance—to make that analysis meaningful, and the categories are unbalanced: in contrast to that 5% of students at the advanced level, about 77% of fourth-graders scored below "proficient" in math in 2000.[11] The benefit of using 10th and 90th percentiles instead is that if NAEP scores are rising equally across all achievement levels, we would expect scores at these two points to behave about the same. Not so with the NAEP achievement-level categories.

The data consist of NAEP reading and math scores for fourth and eighth grades, producing four grade-subject combinations. The NCLB time periods are defined by the last administration of NAEP prior to the law's passage and signing. That is 2000 for fourth-grade math and reading and eighth-grade math. For these three subject-grade combinations, then, the 2000 NAEP serves as the dividing line between pre- and post-NCLB periods in the analysis—that is, as the starting point in the 2000–2007 NCLB-era data, and as the endpoint in the 1990–2000 pre-NCLB data. Eighth-grade reading was not given in 2000, but was given in 1998 and 2002. For that subject-grade combination, 1990–2002 constitutes the pre-NCLB period and 2002–2007 the NCLB-era data.

P-values and standard errors for all of the data reported in the paper appear in tables in appendices A and B, respectively. Please note that the NAEP sample is so large (more than 160,000 students) that even changes of one or two points in a NAEP score—or mean differences of three or four percentage points in a descriptive statistic—can be statistically significant, although perhaps not significant in the real world. In the description below, any value that is described as "large" or "significant" meets significance tests of $p<.05$.

QUESTION 1:
What has happened to the national NAEP scores of high and low achievers since the advent of NCLB?

The four graphs in figure 1 show the NAEP scores of high- and low-achieving students from 2000 to 2007. The graphs on the left, figures 1a and 1b, display scores for fourth grade; those on the right, figures 1c and 1d, display scores for eighth grade. In fourth grade, both high and low achievers made large gains in math (figure 1a). Scores at the 90th percentile rose from 264 to 274, a gain of ten points. Scores at the 10th percentile rose a whopping eighteen points, from 183 to 201. Both gains are statistically significant at $p<.001$. For a more meaningful measure of the magnitude of such gains, a ballpark estimate is that one year of learning is equal to about eleven NAEP points. A gain of eighteen points at the 10th percentile is equal to more than one and a half years of learning, an increase that any teacher or parent of a low-achieving student would surely notice and applaud.

The 2000 tests were the last NAEPs administered before NCLB was proposed, debated in Congress, and signed into law, and 2003 brought the first NAEP test given in math after NCLB went into effect. As figure 1a reveals, the biggest leap in math scores took place from 2000 to 2003. For both low and high achievers, the bulk of the gains of the NCLB era were attained in the very first interval of NAEP testing—from 2000 to 2003. The achievement gap between high and low achievers narrowed immediately after NCLB was passed, but then stabilized.

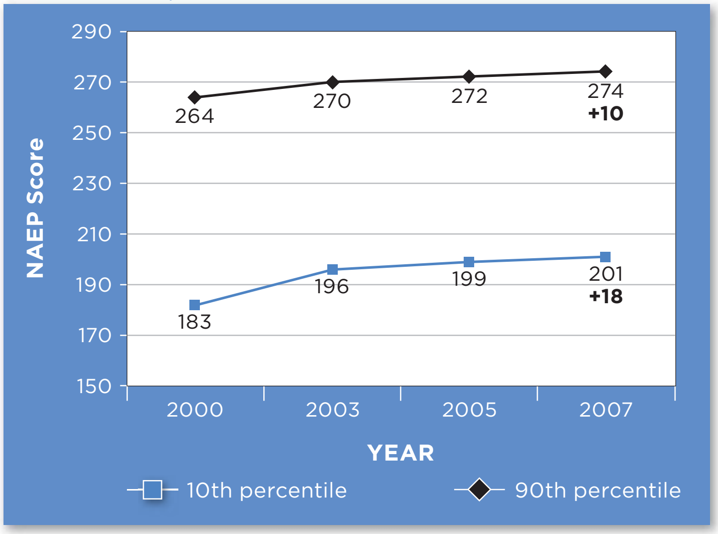
In fourth-grade reading, the sixteen-point gain by low achievers stands out as impressive (see figure 1b). High achievers' scores have remained flat, however. As with math, most of the action in reading scores took place in the initial years. A pop upward of twelve points occurred in low achievers' scores from 2000 to 2002, compared to a one-point gain by high achievers. Over the entire era of NCLB, the gap between the two groups contracted by thirteen scale score points, more than a year's worth of learning.

The eighth-grade scores do not tell a straightforward story. They differ by subject. Math scores follow the same pattern as fourth-grade scores—a pop in low achievers' scores during NCLB's infancy (though not as large as that for fourth-graders), leading to narrowing of the achievement gap, and then similar growth by both low and high achievers in subsequent years (see figure 1c). But eighth-grade reading diverges from this pattern (see figure 1d). From 2002 to 2003, scores at the 90th percentile increased by a point, while the scores at the 10th percentile declined four points, from 219 to 215. The achievement gap widened. From 2003 to 2007, scores for both groups barely budged, with low achievers gaining a point and high achievers losing a point. Over all, unlike the other three grade-subject combinations, eighth-grade reading evidences no progress at the 10th percentile during the NCLB era.

Why is eighth-grade reading an outlier? Note that it has a different baseline year (2002) than the other grade-subject combinations in the analysis because no eighth-grade reading test was given in 2000. Any gains between 2000 and 2002, which are quite large for the other three grade-subject combinations, therefore go undetected. The prior NAEP test in eighth-grade reading was in 1998. From 1998 to 2002, eighth-grade reading did experience a jump in scores, and, interestingly, the 10th percentile gained more than the 90th percentile. The unique nature of eighth-grade NAEP scores in reading should be kept in mind for the remainder of the discussion.

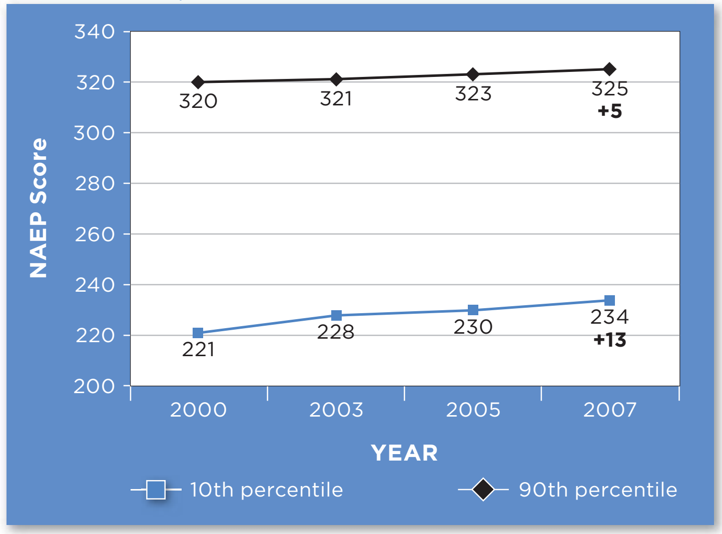
Another important consideration concerning time intervals should also now be apparent from examining the NAEP data. Three grade-subject combinations exhibit a consistent

Figure 1a—Math 4th Grade NAEP Scores, 2000-2007 (90th and 10th percentiles)



Note: *National means: 2000=225, 2007=241, a change of +16*
Source: Main NAEP data explorer, National Public sample

Figure 1c—Math 8th Grade NAEP Scores, 2000-2007 (90th and 10th percentiles)

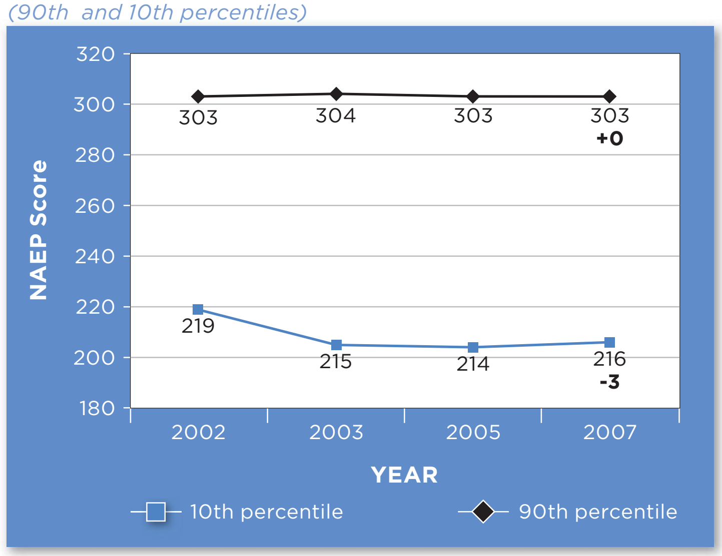


Note: *National means: 2000=225, 2007=241, a change of +16*
Source: Main NAEP data explorer, National Public sample

Figure 1b—Reading 4th Grade NAEP Scores, 2000-2007 (90th and 10th percentiles)

Note: *National means: 2000= 215, 2007=222, a change of +7*
Source: Main NAEP data explorer, National Public sample

Figure 1d—Reading 8th Grade NAEP Scores, 2002-2007 (90th and 10th percentiles)



Note: *National means: 2002= 265 and 2007= 264, a change of -1*
Source: Main NAEP data explorer, National Public sample

pattern, a straightforward story of narrowing gaps during the NCLB era—mostly the result of sharp gains by low-achieving students from 2000 to 2002 or from 2000 to 2003. But whether these years belong in the NCLB era is debatable. The starting point matters. Using the NAEP test immediately before NCLB's passage as a baseline, as this study does, includes growth that may have nothing to do with NCLB. Selecting a later date—2003, for example—and arguing that the act's accountability provisions could not have been implemented before then would lead to the conclusion that growth was much less during the NCLB era (although still statistically significant, as shown in appendix A), and that the gaps between low and high achievers were essentially unchanged. But it would also omit influence that NCLB may have had on NAEP scores during the debate and early implementation of the legislation.

Neal and Schanzenbach provide an example. In the fall of 2001, "with the passage of NCLB looming on the horizon," the state of Illinois placed hundreds of schools on a watch list and declared that future state testing would be high stakes.[12] If such actions influenced educators' behavior and students' test scores, an "NCLB effect" may have been registered in 2002. The bottom line is that there is no clear boundary between pre- and post-NCLB periods and no perfect way to delineate the NCLB era using the NAEP test years. Critics and defenders of NCLB alike can (and do) exploit this ambiguity to their advantage. The fairest approach is to point out the large gains in NAEP scores in the period around 1998–2003 and acknowledge that NCLB's association with these gains is unknown.

Let's turn now to examining NAEP scores from the 1990s to see if the trends for 2000–2007 were evident in the previous decade.

QUESTION 2:
What were the trends in NAEP scores of high and low achievers before NCLB?

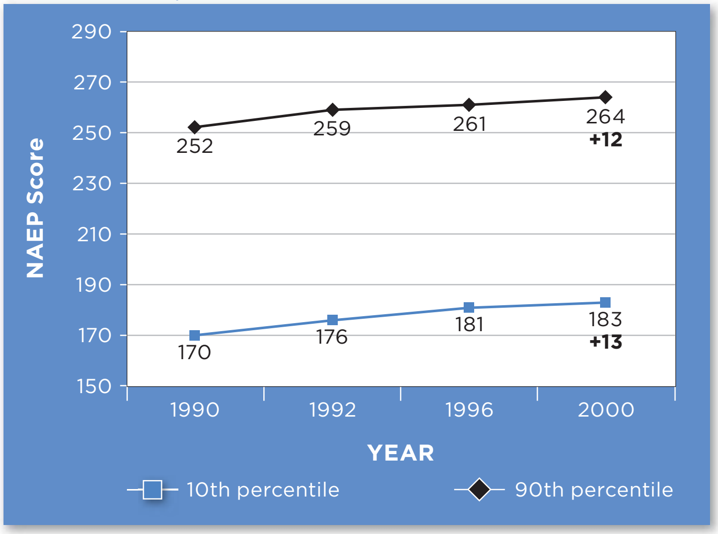
The four graphs of figure 2 display NAEP scores for the 1990s. NAEP testing in the two subjects began in different years: math testing in 1990 and reading in 1992. As figure 2a shows, both high and low achievers in fourth grade made strong gains during the decade. High achievers' scores increased from 252 to 264, a gain of twelve points. Low achievers gained thirteen points, going from 170 to 183. Both gains represent more than a year's worth of learning. The gap between the 10th and 90th percentiles remained essentially unchanged in fourth-grade math.

Fourth-graders as a whole lost ground in reading (see figure 2b). Scores at the 10th percentile fell sharply from 168 to 157, with a large loss from 1992 to 1994. High achievers' reading scores remained flat, ticking up a single point over the entire decade. The gap between high and low achievers expanded in the 1990s due to the declining scores of students at the bottom of the achievement distribution.

The achievement gap also widened in eighth-grade math but for a different reason (see figure 2c). Scores of high achievers moved from 307 to 320, a gain of thirteen points. Low achievers made gains, but not as large—seven points. All boats were rising in eighth-grade math, but the boats at the 90th percentile rose more than those at the 10th percentile. The gap did narrow in eighth-grade reading (see figure 2d). Scores at the 10th percentile rose eight points, in contrast to a one-point gain at the 90th percentile. Thus, math and reading present opposite patterns in eighth grade but, as noted above, the unique time interval for eighth-grade reading scores makes those data difficult to interpret.

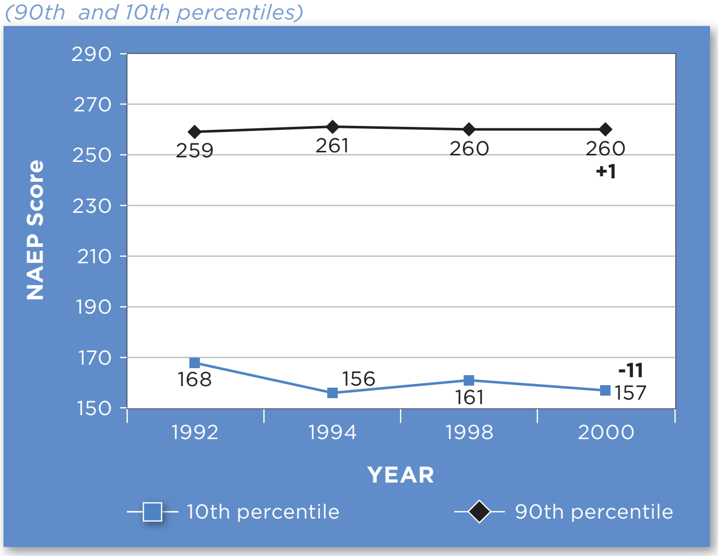
In sum, the 1990s present a mixed picture. The NAEP score gap between high and low achievers widened in fourth-grade

***Figure 2a**—Math 4th Grade NAEP Scores, 1990-2000*
(90th and 10th percentiles)



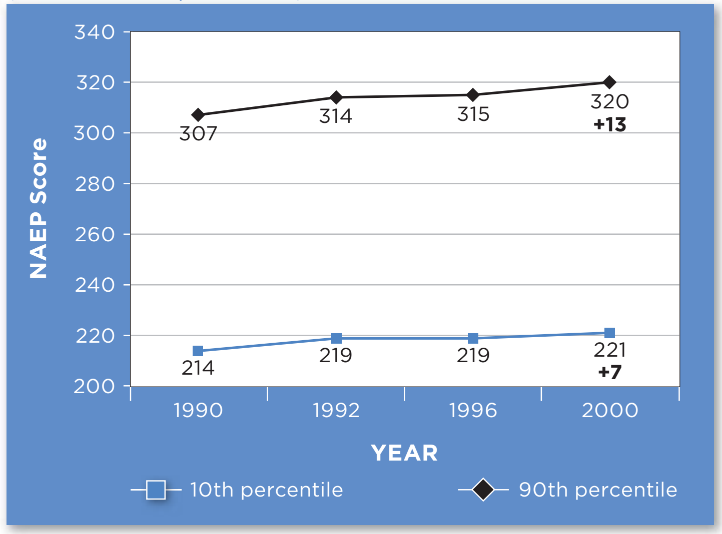
***Note:** National means: 1990= 213, 2000 = 225, a change of +12*
Source: Main NAEP data explorer, National Public sample

***Figure 2b**—Reading 4th Grade NAEP Scores, 1992-2000*
(90th and 10th percentiles)



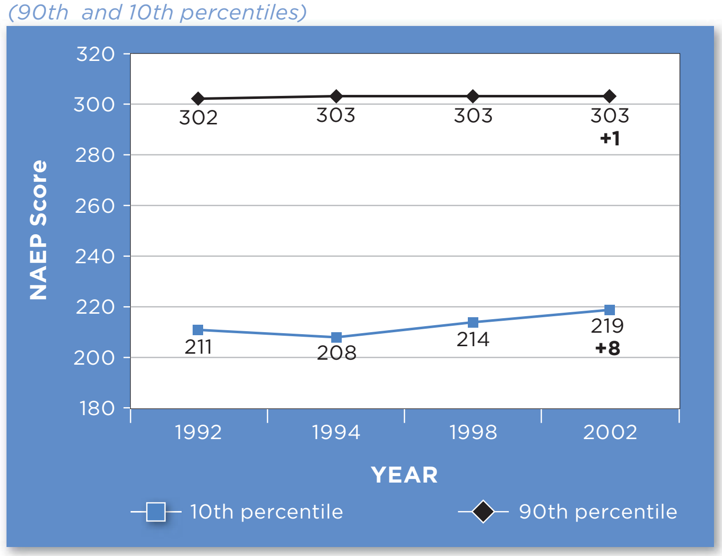
***Note:** National means: 1992 = 217, 2000= 215, a change of -2*
Source: Main NAEP data explorer, National Public sample

***Figure 2c**—Math 8th Grade NAEP Scores, 1990-2000*
(90th and 10th percentiles)



***Note:** National means: 1990= 263 and 2000= 274, a change of +11*
Source: Main NAEP data explorer, National Public sample

***Figure 2d**—Reading 8th Grade NAEP Scores, 1992-2002*
(90th and 10th percentiles)



***Note:** 1992=260 and 2002= 265, a change of +5*
Source: Main NAEP data explorer, National Public sample

Table 1—Annual Gains by 10th percentile pre- and post-NCLB

Grade/Subject	Pre-NCLB	Post-NCLB
Grade 4–Math	1.3	2.6
Grade 4–Reading	-1.4	2.3
Grade 8–Math	0.7	1.9
Grade 4–Reading	0.8	-0.6
Average of grade/subject combinations	0.35	1.55

Average annual gains found by dividing gain over entire interval by number of years in interval. All gains are measured in NAEP scale score points.

Table 2—Annual Gains by 90th percentile pre- and post-NCLB

Grade/Subject	Pre-NCLB	Post-NCLB
Grade 4–Math	1.2	1.4
Grade 4–Reading	0.1	0.4
Grade 8–Math	1.3	0.7
Grade 4–Reading	0.1	0.0
Average of grade/subject combinations	0.675	0.625

Average annual gains found by dividing gain over entire interval by number of years in interval. All gains are measured in NAEP scale score points.

reading and eighth-grade math, but for different reasons. The gap contracted in fourth-grade math and eighth-grade reading, again for different reasons. High achievers generally fared better than low achievers during the 1990s; however, the weak performance of 10th percentile fourth-graders in reading unduly influences this conclusion. Without that steep decline, the conclusion would be that the two groups performed about the same, with both making solid gains.

How do the pre- and post-NCLB periods compare? Tables 1 and 2 report the average annual change in test scores. The changes are expressed in annual increments because the test intervals of the grade-subject combinations span different numbers of years. Table 1 shows changes in test scores for 10th percentile students in the pre- and post-NCLB periods. Table 2 offers the same comparison for 90th percentile students.

The major contrast before and after NCLB occurred in scores at the 10th percentile (see table 1). Low achievers made significant strides on NAEP after 2000. The gains of low achievers in fourth-grade math doubled from the pre-NCLB period (when there was an average annual gain of 1.3 points) to the post-NCLB period (2.6 points). In fourth-grade reading, low achievers lost ground before NCLB (average annual loss of 1.4 points) but accomplished healthy gains after NCLB (2.3 points). The gain in eighth-grade math rose from .7 points per year to 1.9 points per year. And eighth-grade reading exhibits a pattern different from the other grade-subject combinations, showing gains in the pre-NCLB period (average increase of 0.8 points per year) offset by losses during the post-NCLB period (average decline of 0.6 points per year).

For the 90th percentile students, the differences between the two eras' NAEP scores are less pronounced (see table 2). Big gains in fourth-grade math before NCLB (1.2 points per year) continued into the post-NCLB period (1.4 points per year). Trivial gains in fourth-grade reading in the pre-NCLB years were matched by small gains after NCLB. A robust gain of 1.3 points per year in eighth-grade math before NCLB slowed to an average annual gain of 0.7 points during the NCLB era. Scores in eighth-grade reading were flat both before and after NCLB. Overall, growth at the 90th percentile changed very little in the pre- and post-NCLB eras, averaging 0.675 points per year across the four grade-subject combinations in the 1990s and 0.625 after 2000. Growth at the 10th percentile, on the other hand, has averaged 1.55 points per year during the NCLB era, a marked acceleration from the 0.35 points per year in the 1990s. The accelerating growth at the bottom of the achievement distribution is driving the narrowing of the achievement gap.

Let's sum up the data on questions 1 and 2. The national NAEP data support three findings: first, the achievement gap between high and low achievers narrowed during the NCLB era (2000–2007); second, the narrowing of the gap was not taking place immediately prior to NCLB (1990–2000); and third, the narrowing of the gap during the NCLB era is largely due to a significant improvement in the performance of low achievers and smaller gains by high achievers. It is important to stress again that these patterns in NAEP data only indicate correlation and cannot be tied causally to NCLB. But they do confirm the Springer study's finding that NCLB-style accountability is associated with increases in achievement at the bottom of the distribution without declines in achievement at the top.

Holding schools accountable for changes in test scores was not an invention of NCLB. Similar accountability systems were in place in many states in the 1990s. They, too, emphasized boosting the achievement of students at the bottom of the distribution. Maybe, then, accountability in general rather than NCLB accountability in particular is associated with rising scores among low-achieving students. The states present a natural experiment on the question. Some states had accountability systems in the 1990s and some states did not. Examining state NAEP data will allow us to compare them.

QUESTION 3:
Is it NCLB accountability or accountability in general that is associated with contraction of the achievement gap?

NAEP draws on different samples of students to produce national and state scores. This practice provides a way to confirm or reject the trends reported above for national NAEP data. We weighted the gains by population so that a large state counts for more than a small state. Table 3 shows the mean scale score gains at the 10th and 90th percentiles for states that participated in NAEP from 2000 to 2007 (participation was voluntary until 2003). At the beginning of the decade, state math and reading tests were given in different years—2000 for math and 2002 for reading. The statistic for the group of students making the most progress—either the 10th or 90th percentile—is shaded in each row.

In a trend consistent with national NAEP data, low-achieving students made greater academic strides than 90th percentile students on state NAEP tests and narrowed the gap separating the two groups. In fourth-grade math, low achievers notched a 15.5-point gain compared to a 12.8-point gain among high achievers. In fourth-grade reading, low achievers gained 3.6 points versus 1.8 for high achievers. In eighth-grade math, the 10.5-point gain by low achievers outpaced the gain of 8.4 points by high achievers. The general pattern is one of all boats rising; but the boats at the 10th percentile rose more than those at the 90th percentile. Again, eighth-grade reading diverges from the general pattern, with the 90th percentile showing a tiny gain (0.1 point) and low achievers a 2.0-point loss.

For the analysis of pre-NCLB data, we categorized states as having accountability or nonaccountability policies in the 1990s using the coding scheme of Martin Carnoy and Susanna Loeb.[13] They classified as "accountability states" those states with systems that rewarded or sanctioned schools based on test scores. The sanctions of NCLB apply to schools with grades 3–8, which are also the grades of interest in the current study, so if a state's accountability system did not apply to grades 3–8, we re-classified the state as a "nonaccountability" state. In table 4, data are presented for the NAEP testing interval immediately prior to NCLB—1996 to 2000 in math and 1998 to 2002 in reading. Going back earlier in the 1990s would severely diminish the number of states in the analysis since not all states participated in NAEP.[14]

Two questions of interest: Did low achievers gain more than high achievers? And did they gain more in accountability states than in nonaccountability states? The statistic for the group of students making the most progress—either the 10th or 90th percentile—is shaded in each row of table 4. First look at the figures for "overall." The picture is mixed. Low achievers did gain more than high achievers in both subjects at fourth grade—4.5 versus 1.8 points in math and 8.2 versus 2.3 points in reading. But high achievers did better than low achievers in both subjects at eighth grade—a gain of 2.3 points versus a 0.2 loss in math, and a gain of 1.6 points versus a 0.3 gain in reading. So in the NAEP testing period immediately preceding NCLB, the achievement gaps contracted in fourth grade but widened in eighth grade.

The comparison of accountability systems is more decisive. Examine the change in gap statistics for both regimes. Negative values indicate a narrowing gap and positive values a widening gap. For three of the grade-subject combinations, the achievement gap in states with accountability systems improved compared to nonaccountability states. In fourth-grade math, the gap narrowed by 4.1 points in accountability states compared to a widening of 0.6 points in nonaccountability states. In fourth-grade reading, the gap narrowed by 6.6 points in accountability states versus 4.3 points in nonaccountability states. In eighth-grade math,

Table 3—Comparing 90th and 10th Percentile Gains using state NAEP Data. POST-NCLB

Grade/Subject	90th	10th
Grade 4–Math (2000-2007) n=41	12.8	15.5
Grade 4–Reading (2000-2007) n=44	1.8	3.6
Grade 8–Math (2000-2007) n=40	8.4	10.5
Grade 8–Reading (2000-2007) n=42	0.1	-2.0

Note—*All data are in scale score points. Source: Author's calculations from main NAEP data explorer, State NAEP sample.*

the achievement gap expanded in accountability states by 2.2 points but expanded even more (3.4 points) in nonaccountability states. The outlier is eighth-grade reading. The gap expanded by 1.7 points in accountability states and remained unchanged in nonaccountability states.

State NAEP data from the 1990s bolster the theory that accountability systems in general are related to narrower achievement gaps. States that practiced test-based accountability in the 1990s evidence trends in test score gaps that foreshadow what would take place in the NCLB era. But a few wrinkles in the state data from the 1990s must be noted. In the eighth grade, the gap expanded in math, albeit less in accountability states than in nonaccountability states. This is different from the pattern uncovered for the NCLB era, in which the gap in eighth-grade math shrank. And in eighth-grade reading, the constant outlier in these NAEP data, the gap expanded in accountability states and stayed the same in nonaccountability states.

Let's take stock. America's high-achieving students do not appear to have been harmed during the reign of accountability systems—either in the NCLB era or in the era of exclusively state-initiated systems that predate NCLB—though they haven't been helped much, either. The concern about a Robin Hood effect, in which students at the bottom of the achievement distribution make gains at the expense of high achievers, is not substantiated by NAEP data. High achievers' test scores have been rising at a steady, slow pace since 1990. Low achievers' test scores have also been rising, but the pace of those gains increased dramatically sometime between 1998 and 2002—and sooner in states with accountability systems. If the larger gains at the bottom of the achievement distribution are associated with the incentives of accountability systems, this trend suggests a missed opportunity to promote achievement among high achievers.

Table 4—Comparing 90th and 10th Percentile Gains using state NAEP Data. PRE-NCLB

1996-2000 GRADE 4 – MATH	90th	10th	Change in Gap
Accountability | n=16	1.6	5.7	-4.1
Non-accountability | n=20	2.5	1.9	0.6
Overall | n=36	1.8	4.5	-2.7

1998-2002 GRADE 4 – READING	90th	10th	Change in Gap
Accountability | n=16	2.2	8.8	-6.6
Non-accountability | n=21	2.6	6.9	-4.3
Overall | n=37	2.3	8.2	-5.9

1996-2000 GRADE 8 – MATH	90th	10th	Change in Gap
Accountability | n=15	2.5	0.3	2.2
Non-accountability | n=19	2.1	-1.3	3.4
Overall | n=34	2.3	-0.2	2.5

1998-2002 GRADE 8 – READING	90th	10th	Change in Gap
Accountability | n=16	1.5	-0.2	1.7
Non-accountability | n=18	1.9	1.9	0.0
Overall | n=34	1.6	0.3	1.9

***Note**—All data are in scale score points.*
Source: Author's calculations from main NAEP data explorer, State NAEP sample.

QUESTION 4:
Who are America's high achieving students?

The emphasis on closing achievement gaps between high and low achievers places a spotlight on struggling students. Popular media, academic researchers, and public policy devotes considerable resources to students having trouble at school. Often left out of discussions of achievement gaps are high achievers. They are America's best students. What do we know about them?

This section presents a profile of high-achieving students in the United States. The sample of eighth-graders scoring at the 90th percentile or above on NAEP represents about 380,000 pupils. What do we know about them in terms of their demographic characteristics, their schools, and their teachers? We sifted through the restricted-use files of the 2005 NAEP, specifically those pertaining to the eighth-grade math test. Data from the NAEP reading test or another grade might produce different results.[15] Appendix C provides the sources for the independent variables in this question.

THE TYPICAL HIGH ACHIEVER

The typical student scoring at the 90th percentile on the eighth-grade math NAEP comes from a more privileged socioeconomic background than the typical American student (see table 5). Only 10.2% qualify for free or reduced price meals, compared to 36.1% of eighth-graders nationwide and 66.5% of students scoring at the 10th percentile. This means that high achievers are only one-sixth as likely to be eligible for the free or reduced price meals program—a proxy for family income—as low achievers. High achievers also differ from other students in their racial and ethnic backgrounds. More than four out of five (81.5%) of them are white, 2.6% are black, and 4.4% are Hispanic.[16] Among eighth-graders nationwide, 61.1% are white, 16.1% black, and 16.2% Hispanic. The three racial/ethnic groups are fairly evenly represented among low achievers—28.4% white, 36.9% black, and 29.8% Hispanic. As a rule of thumb, blacks and Hispanics are about twice the proportion of low achieving students that one would expect based on the composition of 8th grade students as a whole—and one-fifth to one-fourth of the expected proportion of high achievers.

For several decades, research has identified mothers' education as one of the strongest correlates of family background to student achievement.[17] Nearly two-thirds (64.4%) of high-achieving students have mothers who graduated from college. This is significantly higher than the national average (36.9%) and more than three times the rate for low-achieving students (19.6%). In sum, compared to the typical American eighth-grader, high achievers are more likely to come from higher-income homes, more likely to be white, and less likely to be black or Hispanic, and their mothers are more likely to have earned a college degree.

What math do high achievers study in eighth grade? Table 6 shows enrollment in eighth-grade courses. Most high achievers are enrolled in algebra (57.3%), with a significant number taking geometry (11.1%) or algebra II (4.6%). Thus, nearly three-quarters of high achievers, 73.0%, are taking an advanced math course—algebra or beyond. Among eighth-graders nationwide, almost exactly half as many, 36.6%, are enrolled in such courses. For students at the 10th percentile, the figure is a surprisingly high 28.6%.[18] The recent push to enroll eighth-graders in tougher math courses is apparently paying off, extending even to students for whom mathematics is a struggle. High achievers take advanced math classes, to be sure, but a significant number of low achievers are sitting in the same classrooms.

About 18.0% of high achievers are enrolled in lower-level math classes—pre-algebra, general math, or other (e.g., business math, remedial math)—compared to 61.1% of low-

Table 5—Student Characteristics: 90th Percentile and Comparison Groups

	90th Percentile	National Average	10th Percentile
>50% Eligible Free and Reduced Price Meals	10.2	36.1	66.5
White	81.5	61.1	28.4
Black	2.6	16.1	36.9
Hispanic	4.4	16.2	29.8
Mother is College Grad.	64.4	36.9	19.6

Table 6—Course taking in 8th grade math: 90th Percentile and Comparison Groups

	90th Percentile	National Average	10th Percentile
Geometry	11.1	3.8	5.0
Algebra 2	4.6	3.3	6.2
Algebra 1	57.3	29.5	17.4
2 year Algebra	5.5	4.6	4.6
Pre-Algebra	9.4	26.4	19.2
General Math	6.8	24.4	27.1
Other	1.8	4.8	14.8
Integrated Math	2.9	1.3	1.1

achieving students and 55.6% of eighth-graders overall. Note, though, that these are course titles only and may not reflect the actual quality or rigor of the mathematics taught in the courses. A fruitful line of inquiry for future research would be to investigate eighth-grade math courses and describe how math content varies among courses with the same title.[19]

SCHOOLS ATTENDED BY HIGH ACHIEVERS

The characteristics of schools attended by high achievers are shown in tables 7 and 8. High achievers are more likely to attend suburban schools than other eighth-graders. Low-achieving eighth-graders are more likely to attend urban schools and schools with larger enrollments; these larger schools serve about 885 students compared to a national average of 820 for schools that house an eighth grade (see table 8). The schools of high achievers are average in size, serving 815 students. The negative relationship of school size with achievement—driven here by the presence of low achievers in large schools—has led some school reformers to call for reducing the size of schools.[20]

Let's look at the rest of the characteristics of schools displayed in table 8. Like high-achieving students themselves, the schools of high achievers appear socioeconomically advantaged. About one in seven high achievers (14.7%) attends private schools, much larger than the statistic for eighth-graders nationally (8.8%) and for low achievers (3.3%). Only 10.6% of high-achieving students attend high-poverty schools—those in which at least half of the student body qualifies for free or reduced price meals. That compares to 31.6% of all students nationally and 59.1% of students at the 10th percentile. Only 3.8% of high achievers attend schools with half or more of students receiving targeted Title I services. This is about one-eighth of the figure for low-achieving students (29.7%). Overall, high and low-achievers attend schools with dramatically different demographic profiles.

NAEP asks school principals to report how many students are enrolled in an algebra course in their schools and how many students participate in gifted and talented programs. Both questions are important for determining whether schools are offering high-achieving youngsters educational opportunities that meet their unique educational needs. In 2001, Michigan State researchers examined data from the 1995 Trends in International Mathematics and Science Study (TIMSS) and estimated that one-third of schools did not offer eighth-graders an algebra class.[21] This dismal situation has improved. Evidence supplied by principals in response to the NAEP questionnaire shows that 13.1% of eighth-graders nationwide attended schools without an algebra class in 2005, including 9.2% of high achievers, the students who are presumably best prepared for and most in need of such a course.[22] That still represents about 34,000 students, so despite the improvement, the figure suggests a significant neglect of talent. About 16.5% of low achievers attend schools without algebra, but as indicated above, one-third of low achievers say they are enrolled in advanced math courses. Access to such courses does not appear to be too daunting. Ironically, low-achieving students are more likely to attend schools with gifted programs than high achievers. This may be because access to a variety of programs is intertwined with school size, and attending schools with gifted programs is one benefit that low achievers enjoy in attending larger schools. A less benign possibility is that these gifted programs are used as a substitute for algebra courses and other curricular offerings with truly advanced content.

High-achieving students are more likely to attend schools that assign students to math classes on the basis of ability (i.e., tracking). Among students at the 90th percentile, 78.3% attend a school that tracks eighth-grade math, versus 70.9% for the average student and 65.7% among 10th percentile students. This finding is consistent with research on tracking reform conducted in the 1990s. At that time, an anti-tracking movement swept the country; its proponents argued that such sorting of pupils discriminated against poor and minority children by locking them out of advanced classes.[23] Many low-performing schools, especially in urban areas, responded by abandoning tracking and creating classes of students who were presumably heterogeneous in ability. At the middle-school level, de-tracking was especially popular in English and history departments. Math departments vehemently opposed this reform in the 1990s, but as shown here, even they have been subject to it in many schools. About 22% of high-achieving eighth-graders attend schools that do not group students by ability in mathematics.

TEACHERS OF HIGH ACHIEVERS

What can NAEP tell us about the math teachers of high-achieving students? Three findings stand out (see table 9). They tend to be more experienced than teachers of the typical eighth-grader, with an average of 15.2 classroom years under their belts, compared with 13.5 years for the math teacher of the average eighth-grader and 11.8 for teachers of low-achieving students. A similar pattern is found in the odds of being taught by a novice instructor. Low-achieving students are about twice as likely (29.1%) to have a math teacher in the first four years of his or her career as students at the 90th percentile (16.1%).

Teachers of high achievers are slightly more likely to hold a regular teaching certificate (86.6% versus 82.5% for the average student) and to have majored or minored in math in college. Almost two-thirds of the teachers of high-achieving students majored or minored in math (64.2%) compared to less than half of the teachers of low achievers (44.9%).[24] These data are almost certainly driven by the demographic characteristics of schools. A solid body of research documents dramatic differences in the characteristics of teachers in high- and low-poverty schools, ranging from preparation to experience to turnover.[25] As noted above, high achievers tend to be clustered in low-poverty schools.

The third finding about teachers of high achievers is that they are not walled off from the rest of the students in the schools in which they teach. About one in six of the teachers of 90th percentile students (17.1%) also teach a remedial math class, and four in ten teach general math classes (39.5%). This should allay the concern that teachers of high achievers are cloistered from the general school population and unaware of the needs of average students.

This concern relates to tracking. Critics of tracking argue that grouping kids into classes by ability means that the best students get the best teachers, while kids at risk of failing get the worst teachers. The matching of good teachers and students probably happens innocently. It makes sense that schools assign teachers who know the most math to teach advanced math classes, just as it makes sense that good math students take the toughest math courses. Such commonsense practices create a pairing of staff and students that looks inequitable—high achievers taught by teachers with the strongest math backgrounds and low achievers taught by everybody else, including, of course, those who are weak in math. One way to address the imbalance is to ask more strong math teachers to teach at least one general or remedial math class each day. Another is to increase the supply of teachers with rigorous mathematics training—a longer-term and more satisfying solution but also one that is more ambitious.

***Table 7**—School Locale: 90th Percentile and Comparison Groups*

	90th Percentile	National Average	10th Percentile
Urban	27.5	31.3	43.7
Suburban	51.5	43.1	35.7
Rural	21.0	25.6	20.6

***Table 8**—School Characteristics: 90th Percentile and Comparison Groups*

	90th Percentile	National Average	10th Percentile
School Enrollment	815	820	885
Private School Enrollment	14.7	8.8	3.3
>50% Eligible Free and Reduced Price Meals	10.6	31.6	59.1
>50% Title 1	3.8	14.1	29.7
No Kids in Algebra 1	9.2	13.1	16.5
No Kids in Gifted	26.2	22.8	19.5
8th Grade Math Tracked	78.3	70.9	65.7

***Table 9**—Teacher Characteristics: 90th Percentile and Comparison Groups*

	90th Percentile	National Average	10th Percentile
Teacher Experience (yrs.)	15.2	13.5	11.8
0-4 Years Experience	16.1	22.5	29.1
Regular Teaching Cert.	86.6	82.5	75.8
Major/Minor in Math	64.2	55.8	44.9
Teaches Remedial Math	17.1	24.5	38.3
Teaches General Math	39.5	51.0	57.7

A CLOSER LOOK: HIGH-ACHIEVING STUDENTS FROM THREE NCLB SUBGROUPS

Within the population of high achievers are students targeted by NCLB for special attention. Recall that, in the effort to leave no child behind, NCLB requires schools to break out the test scores of subgroups of children who historically perform below average on tests of academic achievement. What about kids within these subgroups who nonetheless score well above average? From the pool of students scoring at the 90th percentile and above on NAEP, we selected three of these subgroups for special scrutiny—students who are black, Hispanic, and eligible for free or reduced price meals. About 14.0% of high achievers are members of one of these three subgroups, representing approximately 53,000 eighth-graders. They are not being left behind; rather, they are outdistancing their peers in learning. What do the NAEP data tell us about them?

Table 10 displays the socioeconomic characteristics of this group of NCLB high achievers (hereafter called NCLB-HA). Most students in this group come from a lower-income family. Seven out of ten (70.5%) qualify for free or reduced price meals, almost twice the national average. In terms of racial and ethnic backgrounds, the NCLB-HA students are white (39.6%), black (17.8%), and Hispanic (30.5%). The mothers of NCLB-HA students are much more likely to have graduated from college (41.1%) than the mothers of low achievers (19.6%). Indeed, the mothers of NCLB-HA students are more likely to be college grads than are the mothers of average students (36.9%).

The math coursework of NCLB-HA students is somewhat less challenging than that of other 90th percentile students (see table 11). About 64.2% are taking algebra or beyond in eighth grade, nine percentage points less than for the 90th percentile group as a whole. Enrollment by NCLB-HA students in general math and pre-algebra (23.9%) exceeds that of all high achievers (16.2%). These less rigorous courses seem to be drawing students who are capable of handling more advanced mathematics in eighth grade. Do not forget that NCLB-HA students score at the 90th percentile on NAEP—they differ from other high achievers only in race, ethnicity, or family income.[26]

School characteristics for NCLB-HA students are displayed in tables 12 and 13. Table 12 confirms that these students are more likely to attend schools in urban areas (39.0%) compared to other 90th percentile students (27.5%). Indeed, the schools serving NCLB-HA students look more like schools serving 10th percentile students than schools for those at the upper end of the achievement distribution. NCLB-HA students attend larger schools (863 students versus 815 students) and are much less likely to attend private schools than the typical high achiever (see table 13). Features of the large, urban public school carry over into the remaining data in table 13. The schools of NCLB-HA students enroll more youngsters eligible for free or reduced price meals and targeted Title I services than the average school of high achievers.

About one in seven NCLB-HA students (13.3%) attends a school without an algebra class. Interestingly, the percentage of NCLB-HA students attending schools with tracking (71.3%) resembles the national average (70.9%), not the figure for other high achievers (78.3%). These statistics underscore the impact of tracking reform on urban schools. High achievers who are poor, black, or Hispanic are more likely to attend schools that shun tracking than are high-achieving students who are white, come from higher-income homes, or attend suburban schools. To the extent that heterogeneously grouped math classes hold back students who excel at mathematics—and there is some evidence that they do—this limitation falls disproportionately on NCLB-HA students.[27]

Are NCLB-HA students shortchanged on teacher quality? They do not appear to be according to the measures available in NAEP (see table 14). In years of experience, percentage of new teachers, and rates of standard certification, the differences between teachers of NCLB-HA students and high achievers as a whole are not statistically significant (at $p<.05$). Teachers of NCLB-HA students have more experience and higher rates of standard certification than the teacher of the typical American eighth-grader. Moreover, NCLB-HA students are just as likely as other high achievers to have math teachers who majored or minored in the subject in college (64.5% versus 64.2%) and significantly more likely to have such teachers than the average student nationwide (55.8%).

***Table 10**—Student Characteristics: NCLB-HA and Comparison Groups*

	NCLB-HA	90th Percentile	National Average	10th Percentile
>50% Eligible Free and Reduced Price Meals	70.5	10.2	36.1	66.5
White	39.6	81.5	61.1	28.4
Black	17.8	2.6	16.1	36.9
Hispanic	30.5	4.4	16.2	29.8
Mother is College Grad.	41.1	64.4	36.9	19.6

***Table 11**—Course taking in 8th grade math: NCLB-HA and Comparison Groups*

	NCLB-HA	90th Percentile	National Average	10th Percentile
Geometry	8.6	11.1	3.8	5.0
Algebra 2	3.9	4.6	3.3	6.2
Algebra 1	51.7	57.3	29.5	17.4
2 year Algebra	5.6	5.5	4.6	4.6
Pre-Algebra	13.1	9.4	26.4	19.2
General Math	10.8	6.8	24.4	27.1
Other	2.5	1.8	4.8	14.8
Integrated Math	2.9	2.9	1.3	1.1

Table 12—School Locale: NCLB-HA and Comparison Groups

	NCLB-HA	90th Percentile	National Average	10th Percentile
Urban	39.0	27.5	31.3	43.7
Suburban	40.2	51.5	43.1	35.7
Rural	20.9	21.0	25.6	20.6

Table 13—School Characteristics: NCLB-HA and Comparison Groups

	NCLB-HA	90th Percentile	National Average	10th Percentile
School Enrollment	863	815	819.7	885
Private School Enrollment	8.6	14.7	8.8	3.3
>50% Eligible Free and Reduced Price Meals	33.3	10.6	31.6	59.1
>50% Title 1	13.8	3.8	14.1	29.7
No Kids in Algebra 1	13.3	9.2	13.1	16.5
No Kids in Gifted	20.1	26.2	22.8	19.5
8th Grade Math Tracked	71.3	78.3	70.9	65.7

Table 14—Teacher Characteristics: NCLB-HA and Comparison Groups

	NCLB-HA	90th Percentile	National Average	10th Percentile
Teacher Experience (yrs.)	14.3	15.2	13.5	11.8
0-4 Years Experience	20.3	16.1	22.5	29.1
Regular Teaching Cert.	84.2	86.6	82.5	75.8
Major/Minor in Math	64.5	64.2	55.8	44.9
Teaches Remedial Math	20.4	17.1	24.5	38.3
Teaches General Math	46	39.5	51	57.7

SUMMARY AND CONCLUSION

Concerns have been raised about how high-achieving students may be affected by accountability systems, including NCLB. Has the emphasis on getting struggling students over a low academic bar diminished the quality of education for students who excel academically? The NAEP data lead to several conclusions. During the NCLB era, achievement gaps between high- and low-achieving students have narrowed. Both high and low achievers have made test score gains since the federal government debated and implemented NCLB—though not necessarily because of NCLB—but low achievers have gained more. The trend is evident on both national and state NAEP scores.

National NAEP data from the 1990s offer a mixed picture. State NAEP data from the late 1990s also offer a mixed picture, with one important exception: test score changes in states that had accountability systems in place before NCLB look more like the post-NCLB pattern—with all boats rising and low achievers' boats rising more—than those in states that did not have accountability systems. So it appears that accountability systems in general are associated with a similar pattern. The NAEP data trends reported here mirror the state data analyzed by Springer, whose research we looked at briefly above.

A few caveats. To reiterate a point already made, the choice of what year to use as the beginning of the NCLB era affects conclusions about the behavior of test scores during that era. Since the largest gains were accomplished before 2003, starting the era in 2003 will significantly reduce gains made within the era. The second caveat pertains to eighth-grade reading scores. Among the four grade-subject combinations analyzed in the study, it is a constant outlier. The divergence may be due to the different years that the test was administered, but that is only a conjecture, and any conclusions about eighth-grade reading must be made cautiously. Third, the study does not allow for firm conclusions about the effects of NCLB. It is true that the trends reported here are inconsistent with the hypothesis that NCLB's emphasis on low-achieving students somehow cheats high achievers. But the data cannot support or reject claims of causality. Perhaps high achievers would have performed even better if NCLB never existed, or perhaps the trends reported here were caused by other policy interventions or changes in the family or society. NAEP data cannot confirm or rebut such possibilities.

It would be a mistake to allow the narrowing of test score gaps, although an important accomplishment, to overshadow the languid performance trends of high-achieving students. Their test scores are not being harmed during the NCLB era, but they are not flourishing either. Gaps are narrowing because the gains of low-achieving students are outstripping those of high achievers by a factor of two or three to one. The nation has a strong interest in developing the talents of its best students to their fullest to foster the kind of growth at the top end of the achievement distribution that has been occurring at the bottom end. International comparisons of top students around the world invariably show American high-achievers falling short. The data reviewed here offer no indication of that problem being solved anytime soon.[28]

There are several implications to consider from the data on characteristics of high achievers. High achievers possess socioeconomic advantages and more advantaged schools and teachers. Compared to the average pupil—and especially to the typical low-achieving student—they come from higher-income families and their mothers are more educated. They are more likely to attend schools in suburban areas, and their schools are less likely to serve low-income children. They take higher-level math courses and have more experienced teachers, and their math teachers are more likely to have majored or minored in math in college.

Although scoring at equally lofty levels on NAEP, high achievers who come from NCLB-designated groups—black, Hispanic, or low income—evidence a different set of characteristics than their high-achieving peers. These students come from less privileged socioeconomic backgrounds and attend schools with more constraints—larger numbers of poor, urban children and fewer advanced math courses offered. In fact, an eighth-grader who scores at the national average in math is slightly more likely to attend a school with an algebra course than an NCLB-HA student scoring at the 90th percentile. Despite rising scores for high achievers in the NCLB era, these are the students at risk of suffering any lost opportunities stemming from NCLB's incentives.[29]

The math courses offered to NCLB-HA students deserve close scrutiny. As noted above, great progress has been made in providing algebra in most schools. Yet there is room for improvement. About one-quarter of NCLB-HA students (26.4%) are in math classes that precede algebra (pre-algebra, general math, or other) compared to 18.0% of all high-achieving eighth graders. Thousands of excellent math students are not being adequately challenged in the subject—at a time when these students are about to enter high school. The NCLB-HA students have math teachers who appear as qualified to teach advanced courses as the teachers of high achievers as a whole. Granted, the data offer only crude proxies for teacher quality, but they are commonly cited as national and state indicators. Years of teaching experience are similar, and similar percentages hold standard teaching certificates. Similar percentages majored or minored in mathematics. On this last measure—important in preparation to teach algebra, geometry, and advanced algebra—the teachers of NCLB-HA students are more highly qualified than teachers of the average eighth-grader nationwide.

Thus, the lack of advanced math classes appears to be school-based, in the sense that it is a product of school policy or circumstances at schools, not of student or teacher preparation. Some schools may have too few students with the prerequisite skills to handle algebra and therefore cannot fill a single algebra class. The fact that the schools of NCLB-HA students are less likely to group students by ability in math classes could also lead to fewer advanced curricular offerings.

These findings have two sets of policy implications: one directed at schools and districts, the other at policymakers who create accountability systems. If course offerings in math are limited for NCLB-HA students—or anyone else—because of school-based factors, opportunities for taking advanced math need to be opened up that are independent of schools. No eighth-grader who is ready for algebra should be denied access to that subject simply because of the school that he or she attends. The same imperative holds for other advanced math classes. If districts or schools find it impossible to provide these math courses, for whatever reason, then web-based courses should be offered to students who can demonstrate that they are prepared to take them.

The current study joins a growing body of research that suggests that incentives incorporated into accountability systems work about as intended. The key is to get the incentives right. To promote the continued progress of high-achieving students, policymakers should consider creating incentives for schools to boost more students into the upper echelons of achievement.

Here is a modest proposal. Congress should fund an experiment, perhaps as part of the reauthorization of NCLB, that would both add to our understanding of how accountability systems work and create new educational opportunities for gifted disadvantaged youngsters. Schools with large numbers of NCLB-HA students would be invited to participate and randomly assigned to treatment or control groups. Control schools would be subject to standard NCLB provisions. Treatment schools would be eligible for rewards. Rewards would be offered for improving the test scores of high-achieving students, with the reward increasing, perhaps doubling, for gains by students in the NCLB-HA groups. Evaluation could be built into the program so that, after a reasonable period of time, the effects would be assessed and findings released to the public. If the impact turned out to be beneficial, the program could be expanded. Such an experiment might motivate schools to better serve high achievers, improve the image of NCLB by adding carrots to a program with an incentive structure that currently is all sticks, and produce valuable data for policy researchers.

Accountability systems try to improve the education of students who struggle in school, and the preponderance of evidence suggests that they have succeeded in boosting the performance of low achievers. NCLB continues in that tradition. The next generation of accountability in education must build on that accomplishment to maximize the attainments of all students, including America's highest achievers.

NOTES

[1] Sydney Marland Jr., *Education of the Gifted and Talented—Volume 1: Report to the Congress of the United States by the U. S. Commissioner of Education* (Washington, DC: Office of Education, 1971), 6.

[2] Equity and excellence are two major themes of school reform. See Tom Loveless, "Uneasy Allies: The Evolving Relationship of School and State," *Educational Evaluation and Policy Analysis* 20, no.1 (1998): 1–8.

[3] Tom Loveless, "The Peculiar Politics of No Child Left Behind," in *Standards-Based Reform and the Poverty Gap: Lessons for No Child Left Behind*, ed. A. Gamoran (Washington, DC: The Brookings Institution, 2007), 253–85.

[4] Susan Goodkin, "Leave No Gifted Child Behind, *Washington Post*, December 27, 2005, A25.

[5] Derek Neal and Diane Schanzenbach, "Left Behind by Design: Proficiency Counts and Test-Based Accountability," NBER Working Paper No. W13293, 2007.

[6] Matthew Springer, "Accountability Incentives," *Education Next* 8, no. 1 (2008): 74-79.

[7] Randall Reback, "Teaching to the Rating" Journal of Public Economics (2007), doi:10.1016/j.jpubeco.2007.05.003

[8] Ibid, page 3.

[9] P. Clopton, W. Bishop, and D. Klein, "Statewide Mathematics Assessment in Texas," http://www.mathematicallycorrect.com/lonestar.htm (accessed March 14, 2008).

[10] See Lori Shepard et al., *Setting Performance Standards for Student Achievement* (Washington, DC: National Academy of Education, 1993).

[11] Author's calculations from the main NAEP data explorer, http://nces.ed.gov/nationsreportcard/nde/.

[12] Neal and Schanzenbach, "Left Behind," p. 13.

[13] Martin Carnoy and Susanna Loeb, "Does External Accountability Affect Student Outcomes? A Cross-State Analysis," *Educational Evaluation and Policy Analysis* 24, no. 4 (2002): 305–31.

[14] Carnoy and Loeb also took this approach in order to maximize the number of states available for analysis, as did Eric A. Hanushek and Margaret E. Raymond, "High-Stakes Research," *Education Next*, no. 3 (2003): 48-55.

[15] Test scores for Hispanic students, because of potential language barriers affecting reading performance, may especially differ between the math and reading tests.

[16] Black, white, and Hispanic students constitute about 90% of the sample. The 10% of pupils unaccounted for in the discussion are Asian or other racial/ethnic classifications.

[17] This correlation was noted over forty years ago by James S. Coleman et al., *Equality of Educational Opportunity* (Washington, DC: U.S. Government Printing Office, 1966) and has been confirmed by subsequent research.

[18] This does not include 4.6% of students who are enrolled in two-year algebra courses, which are offered to increase the number of students taking algebra by slowing down the pace and stretching the curriculum over two years.

[19] Reba Page, *Lower-Track Classrooms: A Curricular and Cultural Perspective* (New York: Teachers College Press, 1991).

[20] Tom Loveless and Frederick Hess, "Introduction: What Do We Know about School Size and Class Size?" in *Brookings Papers on Education Policy: 2006–2007*, ed. T. Loveless & F. Hess, (Washington, DC: The Brookings Institution Press, 2007), 1-14.

[21] L. S. Cogan, W. H. Schmidt, and D. E. Wiley, "Who Takes What Math and in Which Track? Using TIMSS to Characterize U.S. Students' Eighth-Grade Mathematics Learning Opportunities," *Educational Evaluation and Policy Analysis 23*, no. 4 (Winter 2001): 323–41.

[22] A technical issue: principals were asked not if their schools offered an algebra class, but how many students were taking an algebra class (and geometry and other math classes). The ensuing discussion assumes that schools do not offer an algebra class if the principal reported zero students taking algebra.

[23] Jeannie Oakes [*Keeping Track:How Schools Structure Inequality* (New Haven: Yale University Press, 1985)] was influential in spurring the anti-tracking movement, which continues today. For research on schools responding to tracking reform, see Tom Loveless, *The Tracking Wars: State Reform Meets School Policy* (Washington, DC: The Brookings Institution, 1999).

[24] Tom Loveless, *The 2004 Brown Center Report on American Education: How Well Are Students Learning?* (Washington, DC: The Brookings Institution, 2004).

[25] L. Darling-Hammond, "Teachers and Teaching: Testing Policy Hypotheses from a National Commission Report," *Educational Researcher* 27, no. 1 (1998): 5–15.

[26] Test scores aren't everything in determining whether students are ready for particular classes. Study skills, attendance, and other characteristics may come into play. Moreover, the eighth-grade NAEP may not be the appropriate test for differentiating students who are equipped to take algebra from those who are not.

[27] See Daniel I. Rees, Laura M. Argys, and Dominic J. Brewer, "Tracking in the United States: Descriptive Statistics from NELS," *Economics of Education Review* 15, no. 1 (1996): 83–89. See also Chen-Lin C. Kulik, and James A. Kulik, "Effects of Ability Grouping on Secondary School Students: A Meta-Analysis of Evaluation Findings," *American Educational Research Journal* 19 (1982): 415–28.

[28] Patrick Gonzales, "America's Top-Performing Students in Mathematics and Science: How Do They Compare to Their International Peers?" Paper presented at the American Association for the Advancement of Science (AAAS) 2007 annual conference, San Francisco, CA, February 17, 2007.

[29] We examined the performance of NCLB-HA students on NAEP in 2000 and 2005 and found that they gained about 4.6 scale score points in math, statistically indistinguishable from the gains of the 90th percentile as a whole. The percentage of high achievers falling into one of the NCLB-HA subgroups groups rose from 9.4% to 14.5% during this period, but the share of the subgroups also rose nationally, so this is not unexpected.

APPENDIX A1—*P-Values for Gains at the 90th and 10th Percentiles in the Era of NCLB*

4th Grade Math	2000-2007	2003-2007
90th Percentile	p<.001	p<.001
10th Percentile	p<.001	p<.001

4th Grade Reading	2000-2007	2003-2007
90th Percentile	p<.05	NS
10th Percentile	p<.001	p<.001

8th Grade Math	2000-2007	2003-2007
90th Percentile	p<.001	p<.001
10th Percentile	p<.001	p<.001

8th Grade Reading	2002-2007	2003-2007
90th Percentile	NS	NS
10th Percentile	p<.01	NS

NS=Not statistically significant

APPENDIX A2—*P-Values for Gains at the 90th and 10th Percentiles Pre-NCLB*

4th Grade Math	1990-2000
90th Percentile	p<.001
10th Percentile	p<.001

4th Grade Reading	1992-2000
90th Percentile	NS
10th Percentile	p<.01

8th Grade Math	1990-2000
90th Percentile	p<.001
10th Percentile	p<.01

8th Grade Reading	1992-2000
90th Percentile	NS
10th Percentile	p<.001

NS=Not statistically significant

APPENDIX B—*Descriptive Statistics and Standard Errors*

***Table 5**—Student Characteristics: 90th Percentile and Comparison Groups*

	90th Percentile	National Average	10th Percentile
>50% Eligible Free and Reduced Price Meals	10.2 (.48)	36.1 (.29)	66.5 (.71)
White	81.5 (.60)	61.1 (.31)	28.4 (.56)
Black	2.6 (.25)	16.1 (.23)	36.9 (.71)
Hispanic	4.4 (.29)	16.2 (.22)	29.8 (.76)
Mother is College Grad.	64.4 (.63)	36.9 (.21)	19.6 (.41)

***Table 6**—Math Course Taken in 8th Grade: 90th Percentile and Comparison Groups*

	90th Percentile	National Average	10th Percentile
Geometry	11.1 (.48)	3.8 (.09)	5.0 (.28)
Algebra 2	4.6 (.37)	3.3 (.08)	6.2 (.31)
Algebra 1	57.3 (.74)	29.5 (.20)	17.4 (.52)
2 year Algebra	5.5 (.35)	4.6 (.10)	4.6 (.23)
Pre-Algebra	9.4 (.39)	26.4 (.27)	19.2 (.53)
General Math	6.8 (.33)	24.4 (.26)	27.1 (.56)
Other	1.8 (.19)	4.8 (.07)	14.8 (.35)
Integrated Math	2.9 (.31)	1.3 (.08)	1.1 (.11)

***Table 7**—School Locale: 90th Percentile and Comparison Groups*

	90th Percentile	National Average	10th Percentile
Urban	27.5 (.83)	31.3 (.34)	43.7 (.72)
Suburban	51.5 (.96)	43.1 (.38)	35.7 (.76)
Rural	21.0 (.62)	25.6 (.28)	20.6 (.58)

APPENDIX B—*Descriptive Statistics and Standard Errors (continued)*

***Table 8**—School Characteristics: 90th Percentile and Comparison Groups*

	90th Percentile	National Average	10th Percentile
School Enrollment	814.7 (11.06)	819.7 (7.57)	885 (12.12)
Private School Enrollment	14.7 (.68)	8.8 (.16)	3.3 (.31)
>50% Eligible Free and Reduced Price Meals	10.6 (.56)	31.6 (.50)	59.1 (.92)
>50% Title 1	3.8 (.37)	14.1 (.47)	29.7 (1.10)
No. Kids in Algebra 1	9.2 (.62)	13.1 (.44)	16.5 (.84)
No. Kids in Gifted	26.2 (.84)	22.8 (.55)	19.5 (.79)
8th Grade Math Tracked	78.3 (.94)	70.9 (.61)	65.7 (.96)

***Table 9**—Teacher Characteristics: 90th Percentile and Comparison Groups*

	90th Percentile	National Average	10th Percentile
Teacher Experience (yrs.)	15.2 (.19)	13.5 (.12)	11.8 (.20)
0-4 Years Experience	16.1 (.71)	22.5 (.51)	29.1 (.92)
Regular Teaching Cert.	86.6 (.65)	82.5 (.42)	75.8 (.74)
Major/Minor in Math	64.2 (1.10)	55.8 (.52)	44.9 (.92)
Teaches Remedial Math	17.1 (.79)	24.5 (.53)	38.3 (.92)
Teaches General Math	39.5 (.90)	51.0 (.61)	57.7 (1.01)

APPENDIX B—*Descriptive Statistics and Standard Errors (continued)*

***Table 10**—Student Characteristics: NCLB-HA and Comparison Groups*

	NCLB-HA	90th Percentile	National Average	10th Percentile
>50% Eligible Free and Reduced Price Meals	70.5 (1.52)	10.2 (.48)	36.1 (.29)	66.5 (.71)
White	39.6 (1.76)	81.5 (.60)	61.1 (.31)	28.4 (.56)
Black	17.8 (1.36)	2.6 (.25)	16.1 (.23)	36.9 (.71)
Hispanic	30.5 (1.51)	4.4 (.29)	16.2 (.22)	29.8 (.76)
Mother is College Grad.	41.1 (2.04)	64.4 (.63)	36.9 (.21)	19.6 (.41)

***Table 11**—Math Course Taken in 8th Grade: NCLB-HA and Comparison Groups*

	NCLB-HA	90th Percentile	National Average	10th Percentile
Geometry	8.6 (.83)	11.1 (.48)	3.8 (.09)	5.0 (.28)
Algebra 2	3.9 (.70)	4.6 (.37)	3.3 (.08)	6.2 (.31)
Algebra 1	51.7 (1.67)	57.3 (.74)	29.5 (.20)	17.4 (.52)
2-year Algebra	5.6 (.90)	5.5 (.35)	4.6 (.10)	4.6 (.23)
Pre-Algebra	13.1 (1.41)	9.4 (.39)	26.4 (.27)	19.2 (.53)
General Math	10.8 (1.19)	6.8 (.33)	24.4 (.26)	27.1 (.56)
Other	2.5 (.57)	1.8 (.19)	4.8 (.07)	14.8 (.35)
Integrated Math	2.9 (.58)	2.9 (.31)	1.3 (.08)	1.1 (.11)

***Table 12**—School Locale: NCLB-HA and Comparison Groups*

	NCLB-HA	90th Percentile	National Average	10th Percentile
Urban	39.0 (1.91)	27.5 (.83)	31.3 (.34)	43.7 (.72)
Suburban	40.2 (1.85)	51.5 (.96)	43.1 (.38)	35.7 (.76)
Rural	20.9 (1.30)	21.0 (.62)	25.6 (.28)	20.6 (.58)

APPENDIX B—*Descriptive Statistics and Standard Errors (continued)*

***Table 13**—School Characteristics: NCLB-HA and Comparison Groups*

	NCLB-HA	90th Percentile	National Average	10th Percentile
School Enrollment	862.9 (18.58)	814.7 (11.06)	819.7 (7.57)	885 (12.12)
Private School Enrollment	8.6 (1.45)	14.7 (.68)	8.8 (.16)	3.3 (.31)
>50% Eligible Free and Reduced Price Meals	33.3 (2.16)	10.6 (.56)	31.6 (.50)	59.1 (.92)
>50% Title 1	13.8 (1.39)	3.8 (.37)	14.1 (.47)	29.7 (1.10)
No Kids in Algebra 1	13.3 (1.72)	9.2 (.62)	13.1 (.44)	16.5 (.84)
No Kids in Gifted	20.1 (2.06)	26.2 (.84)	22.8 (.55)	19.5 (.79)
8th Grade Math Tracked	71.3 (1.65)	78.3 (.94)	70.9 (.61)	65.7 (.96)

***Table 14**—Teacher Characteristics: NCLB-HA and Comparison Groups*

	NCLB-HA	90th Percentile	National Average	10th Percentile
Teacher Experience (yrs.)	14.3 (.45)	15.2 (.19)	13.5 (.12)	11.8 (.20)
0-4 Years Experience	20.3 (1.80)	16.1 (.71)	22.5 (.51)	29.1 (.92)
Regular Teaching Cert.	84.2 (1.84)	86.6 (.65)	82.5 (.42)	75.8 (.74)
Major/Minor in Math	64.5 (1.86)	64.2 (1.10)	55.8 (.52)	44.9 (.92)
Teaches Remedial Math	20.4 (1.35)	17.1 (.79)	24.5 (.53)	38.3 (.92)
Teaches General Math	46 (2.24)	39.5 (.90)	51 (.61)	57.7 (1.01)

APPENDIX C—*Sources for Independent Variables in Question 4*

The descriptive variables were taken directly from restricted-use NAEP data files. In some cases NAEP data included collapsed versions of variables that we chose to use. These cases are noted where applicable. We have listed the variable ID along with the variable's source in the student, teacher, or school background questionnaires. Student, school, and teacher background questionnaires from the 2005 NAEP can be retrieved at http://nces.ed.gov/nationsreportcard/bgquest.asp.

STUDENT DEMOGRAPHIC CHARACTERISTICS

1. Eligible Free Lunch (SLUNCH01)—collapsed version of SLNCH05

2.5.3 Eligibility for the Free and Reduced Price Meals Program (SLNCH05)
"Based on available school records for the free/reduced-price lunch component of the Department of Agriculture's National School Lunch Program (http://www.fns.usda.gov/cnd/), students were classified as either: currently eligible, not currently eligible, eligible for reduced-price lunch, not participating, or information not available. The classification refers only to the school year when the assessments were administered (i.e., the 2004–2005 school year) and is not based on eligibility in previous years. If school records were not available, the student was classified as 'Information not available.' If the school did not participate in the program, all students in that school were classified as 'Information not available.'" A. M. Rogers and J. J. Stoeckel, *NAEP 2006 Mathematics, Reading, and Science Restricted-Use Data Files Data Companion*, Mathematics (NCES 2007-485, NCES 2007-486) (Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2007), 34.

2. White, Black, Hispanic (SDRACEM)

2.5.2 Race/Ethnicity (SDRACEM)
"In all NAEP assessments, data about student race/ethnicity is collected from two sources: school records and student self-reports. Before 2002, NAEP used students' self-reports of their race and ethnicity on a background questionnaire as the source of race/ethnicity data. In 2002, it was decided to change the student race/ethnicity variable highlighted in NAEP reports. Starting in 2002, NAEP reports of students' race and ethnicity are based on the school records, with students' self-reports used only if school data are missing. The resulting variable SDRACEM contains a value for every student." A. M. Rogers and J. J. Stoeckel, *NAEP 2006 Mathematics, Reading, and Science Restricted-Use Data Files Data Companion,* Mathematics (NCES 2007-485, NCES 2007-486) (Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2007), 34.

3. Mother's Education B003501
(Section 3, Question 11, Student Background Questionnaire)

How far in school did your mother go?

— She did not finish high school
— She finished high school
— She had some education after high school
— She graduated from college
— I don't know

4. Courses Taken by the 90th Percentile at 8th Grade
M815701 (Section 4, Question 1, Student Background Questionnaire)

What math class are you taking this year?

— Geometry
— Algebra II
— Algebra I (one-year course)
— First year of a two-year Algebra I course
— Second year of a two-year Algebra I course
— Introduction to algebra or pre-algebra
— Basic or general eighth-grade math
— Integrated or sequential math
— Other math class

5. School Locale TOL3
The National Center for Education Statistics (NCES) merged several locale variables from the Common Core of Data with the school-level NAEP variables. We used the three-level variable, TOL3, that collapses locale into urban, suburban, and rural.

SCHOOL CHARACTERISTICS

6. School Enrollment C038101
(Part 1, Question 5, School Background Questionnaire)

What is the current enrollment of your school?

7. Private School Enrollment SCHTYP2
(collapsed version of Question 7 in Part 1, School Background Questionnaire)

What type of school is this? Fill in ovals for all that apply.

— Regular middle or secondary school
— A regular school with a magnet program
— A magnet school or a school with a special program emphasis, e.g., science/math school, performing arts school, talented/gifted school, foreign language immersion school, etc.
— Special education: a school that primarily serves students with disabilities
— Alternative: a school that offers a curriculum designed to provide alternative or nontraditional education, not clearly categorized as regular, special education, or vocational
— Private (independent)
— Private (religiously affiliated)
— Charter school
— Privately run public school
— Other

8. >50% Eligible Free and Reduced Price Meals C051601
(Part 1, Question 11, School Background Questionnaire)

During this school year, about what percentage of students in your school was eligible to receive a free or reduced-price lunch through the National School Lunch Program?

0%	11–25%	51–75%
1–5%	26–34%	76-99%
6–10%	35–50%	100%

9. >50% Eligible Title 1 C051801
(Part 1, Question 13, School Background Questionnaire)

Approximately what percentage of students in your school receives the following services? Fill in one oval on each line. Students who receive more than one service should be counted for each service they receive. Please report the percentage of students who receive each of the following services as of the day you respond to this questionnaire.

a) Targeted Title I Services

None	11–25%	76–90%
1–5%	26–50%	Over 90%
6–10%	51–75%	

10. No Kids in Algebra 1 C052803
(Part 2, Question 3, School Background Questionnaire)

What percentage of eighth-grade students in your school is enrolled in the following mathematics classes? Fill in one oval on each line.

c. Algebra I (one-year course)

None	51–75%
1–10%	76–90%
11–25%	91–100%
26–50%	

11. No Kids in Gifted C044004
(Part 1, Question 13, School Background Questionnaire)

b) Gifted and talented program

None	26–50%
1–5%	51–75%
6–10%	76–90%
11–25%	Over 90%

12. 8th-Gr. Math Tracked C052901
(Part 2, Question 4, School Background Questionnaire)

Are eighth-grade students typically assigned to mathematics classes by ability and/or achievement levels (so that some classes are higher in average ability and/or achievement levels than others)?

—Yes
—No

TEACHER CHARACTERISTICS

13. Teaching Experience (Years) T077101
(Part 1, Question 3, Teacher Background Questionnaire)

Counting this year, how many years have you worked as an elementary or secondary teacher? If less than 4 months total experience, enter "00."

14. 0–4 Years' Experience YRSEXP
NCES collapsed the continuous teaching experience variable into the following categories: 0–4 years, 5–9 years, 10–19 years, 20 years.

15. Regular Teaching Cert. T077201
(Part 1, Question 5, Teacher Background Questionnaire)

What type of teaching certificate do you hold in the state where you currently teach?

— Regular or standard state certificate or advanced professional certificate
— Probationary certificate (the initial certificate issued after satisfying all requirements except the completion of a probationary period)
— Provisional or other type of certificate given to persons who are still participating in what the state calls an "alternative certification program"
— Temporary certificate (requires some additional college coursework and/or student teaching before regular certification can be obtained)
— Emergency certificate or waiver (issued to persons with insufficient teacher preparation who must complete a regular certification program in order to continue teaching)
—No certificate

16. Major/Minor in Math T077310
(Part 1, Question 8, Teacher Background Questionnaire)

Did you have a major, minor, or special emphasis in any of the following subjects as part of your undergraduate coursework? Fill in one oval on each line.

b) Mathematics

— Yes, a Major
— Yes, a Minor or special emphasis
— No

17. Teaches Remedial Math T090801
(Part 1, Question 14, Teacher Background Questionnaire)

Are you teaching the following mathematics courses to eighth-grade students this year? Include honors sections. Fill in one oval on each line.

a) Remedial mathematics

— Yes
— No

18. Teaches General Math T090802
(Part 1, Question 14, Teacher Background Questionnaire)

Are you teaching the following mathematics courses to eighth-grade students this year? Include honors sections. Fill in one oval on each line.

b) General mathematics

— Yes
— No