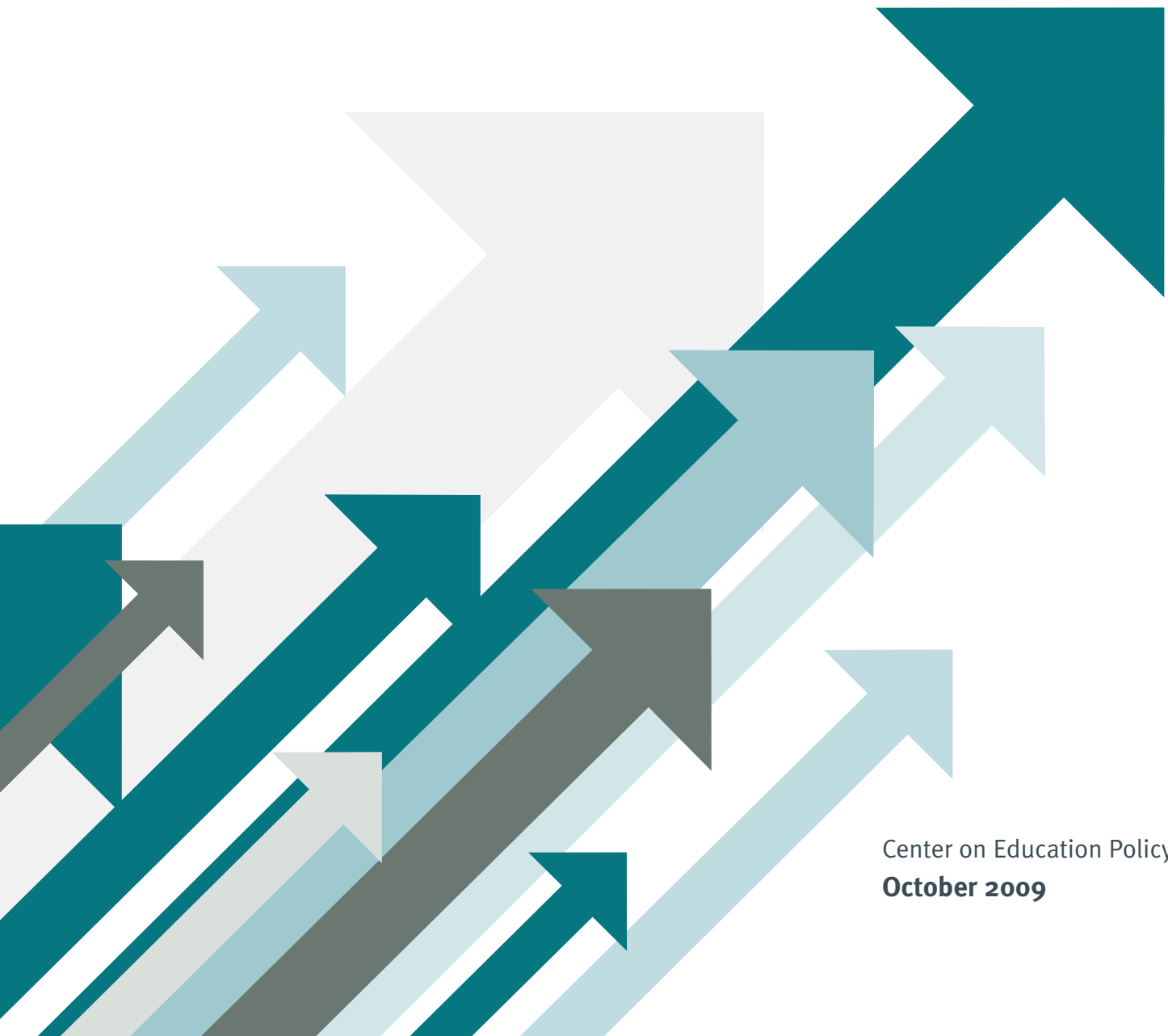


*State Test Score Trends Through 2007-08, Part 3*

# Are Achievement Gaps Closing and Is Achievement Rising for All?



Center on Education Policy  
**October 2009**

*State Test Score Trends Through 2007-08, Part 3*

# Are Achievement Gaps Closing and Is Achievement Rising for All?

# Table of Contents

Introduction .....	1
Main Findings .....	1
Background and Study Methods .....	2
Grade 4 Subgroup Trends at Three Achievement Levels .....	5
Gap Trends Based on Percentages Proficient at Three Grade Levels .....	9
Gap Trends Based on Mean Scores at Three Grade Levels .....	19
Conclusion .....	22
References .....	25

## List of Figures, Tables, and Boxes

Table 1. READING: Number and percentage of states showing various trends for subgroups at three achievement levels on state grade 4 tests .....	7
Table 2. MATH: Number and percentage of states showing various trends for subgroups at three achievement levels on state grade 4 tests .....	8
Table 3. GRADE 4: Number of states showing various trends in percentage proficient gaps on state reading and math tests .....	10
Table 4. MIDDLE SCHOOL: Number of states showing various trends in percentage proficient gaps on state reading and math tests .....	11
Table 5. HIGH SCHOOL: Number of states showing various trends in percentage proficient gaps on state reading and math tests .....	12
Box A. Long-term trends in achievement gaps .....	12
Table 6. Size of achievement gaps in scale scores on the long-term NAEP, 1970s through 2008 .....	12
Table 7. Details on narrowing and widening gaps in percentages proficient .....	15
Table 8. Percentage of gaps narrowing by subgroup, grade level, and subject .....	16
Table 9. Number of states in which the Asian subgroup had a higher percentage proficient than the white subgroup, 2008 .....	18
Table 10. GAPS BY SUBGROUP: Comparison of percentage proficient and mean score measures of achievement gaps .....	21
Table 11. GAPS BY SUBJECT AND GRADE LEVEL: Comparison of percentage proficient and mean score measures of achievement gaps .....	21
Box B. Why gaps change in size depending on where the cut score is set .....	23
Figure A. African American-white achievement gap, NAEP mathematics grade 8, 2000 .....	23
Figure B. Size of gaps in percentages proficient with a cut score at the mean .....	24
Figure C. Size of gaps in percentages proficient with a lower cut score .....	24

# Are Achievement Gaps Closing and Is Achievement Rising for All?

## Introduction

A main goal of the No Child Left Behind Act (NCLB) is to close gaps in test scores between different groups of students while raising achievement for all groups. Of particular concern are the persistent achievement gaps between African American, Latino, and Native American students and their white and Asian counterparts, and between students from low-income families and those who are not low-income. For these gaps to narrow and eventually close, not only must achievement for lower-scoring subgroups increase, but it must go up at a faster rate than for the higher-scoring comparison group. A related question is whether achievement for all subgroups has improved across the scoring scale—at the “advanced” and “basic” levels as well as at the “proficient” level, which receives the most scrutiny under NCLB.

In 2009, the third year of a multiyear study of student achievement, the Center on Education Policy (CEP) did three types of analyses of data broken out by subgroups from the state tests used for NCLB accountability. First, we examined grade 4 state test results at three achievement levels—basic-and-above, proficient-and-above, and advanced. The goal was to see whether students from various racial/ethnic groups and low-income students have made gains at all three achievement levels and whether progress is lagging at any level for specific subgroups. Second, we looked at gaps between subgroups in the percentages of students scoring at or above the proficient level at three grade levels (grade 4, grade 8 in most cases, and a high school grade). Third, we analyzed gaps between subgroups in average test scores at three grade levels. We did these last two analyses to see whether achievement gaps at the elementary, middle, and high school grades have narrowed, widened, or stayed the same since 2002, when NCLB took effect and required states and school districts to pay greater attention to the achievement of subgroups.

This report describes the trends we found for African American, Latino, Native American, Asian, and white students, and for low-income and non-low-income students. Achievement trends for students with disabilities, English language learners, and male and female students will be discussed in future reports in this series.

## Main Findings

Several main findings emerged from our study of subgroup trends:

- **All subgroups showed more gains than declines in grade 4 at all three achievement levels—basic-and-above, proficient-and-above, and advanced.** Trends varied, however, by subject and subgroup. Subgroup results were more positive in math than in reading at all achievement levels. Four-fifths or more of the states with adequate data made gains in math for nearly all subgroups at both the proficient and advanced levels. Progress in math was particularly noteworthy for Latino students scoring at or above the proficient level, where 95% of the states with data improved, and for African Americans scoring at the advanced level, where 91% of the states with data made gains. In reading,

subgroup gains were more common at the proficient level than at the basic or advanced level. A sizeable share of the states with data—between one-fourth and one-third, depending on the subgroup—experienced declines in the percentage of students reaching the advanced level in reading.

- As measured by percentages of students scoring proficient, gaps between subgroups have narrowed in most states at the elementary, middle, and high school levels, although in a notable minority of cases gaps have widened.** In both reading and math, gaps in percentages proficient narrowed in 74% of instances we studied and widened in 23% of instances. For example, the African American-white gap in 4<sup>th</sup> grade reading narrowed in 28 of the 36 states with adequate data, widened in 7 states, and showed no change in 1 state. Across subgroups and states, there was more progress in closing gaps at the elementary and middle school levels than at the high school level. Even with this progress, however, the gaps between subgroups often remained large—upwards of 20 percentage points in many cases.
- Most often gaps narrowed because the achievement of lower-performing subgroups went up rather than because the achievement of higher-performing subgroups went down.** When gaps narrowed, it was usually because both of the subgroups being compared made gains, but the target group (such as Latino students) improved at a greater rate than the comparison group (such as white students). When gaps widened, it was most often because both subgroups made gains but the comparison group improved at a greater rate than the target group.
- Gaps in percentages proficient narrowed more often for the Latino and African American subgroups than for other subgroups.** The Latino-white gap narrowed in 79% of the instances analyzed, and the African American-white gap narrowed in 77% of instances. Results were still positive but less so for the low-income and Native American subgroups. Only a limited number of states, however, had sufficient data to make gap comparisons for Native Americans.
- Although average (mean) scores indicate that gaps have narrowed more often than they have widened, mean scores give a less rosy picture of progress in closing achievement gaps than percentages proficient.** In the states with both mean score and percentage proficient data, gaps narrowed 59% of the time using mean scores, compared with 71% of the time using percentages proficient. Mean score gaps also widened more often than percentage proficient gaps—37% of the time for mean scores versus 24% of the time for percentages proficient. Even so, mean scores for subgroups have risen in most cases, and gaps in mean scores have narrowed in the majority of instances.

## Background and Study Methods

This report is part 3 of CEP’s 2009 series, *State Test Score Trends Through 2007-08*, which looks at student achievement trends since NCLB was enacted in 2002. Part 1 of the series, *Is the Emphasis on Proficiency Shortchanging Higher- and Lower-Achieving Students?* examined trends at the basic, proficient, and advanced levels of achievement for students as a whole (CEP, 2009a). That analysis found that many more states had gains than declines at all three achievement levels. Part 2 of the series, *Is There a Plateau Effect in Test Scores?* concluded that the so-called “plateau” effect—a leveling off of achievement gains after a test has been in place for several years—was neither widespread nor inevitable for the current generation of

state tests (CEP, 2009b). These two reports can be downloaded from CEP's Web site at [www.cep-dc.org](http://www.cep-dc.org).

Achievement gaps between various groups of students have been one focus of our research since CEP began studying student achievement trends in 2007. In last year's report on achievement, we noted that gaps had narrowed in most states, as measured by both state tests and the state-level National Assessment of Educational Progress (NAEP) (CEP, 2008). However, there was only moderate agreement between the two measures. Achievement gap trends on NAEP moved in the same direction as gap trends on state tests 60% of the time in states with sufficient data. Because the state-level NAEP was last administered in 2007 and was covered in our 2008 report, this year's report focuses on state test score trends only.

This report revises and expands on our previous work on subgroup achievement in three main ways. First, it explores a new issue by analyzing trends in grade 4 for subgroups at three achievement levels (basic-and-above, proficient-and-above, and advanced). Second, it includes an additional year (2007-08) of test data, which creates longer trend lines for subgroups. Third, it looks at achievement gaps not only in terms of the percentages of students in each subgroup reaching the proficient level on state tests, but also in terms of mean (average) test scores for each subgroup expressed on the scale for each state's test. These mean scores offer a second measure for gauging whether gaps have narrowed that is not dependent on where a state sets its cut score for proficient performance.

In our analyses of achievement gaps, we compared the performance of "target" subgroups of students (African American, Latino, Native American, or low-income students) with that of a "comparison" group (white students in the case of racial/ethnic minority subgroups, and non-low-income students in the case of the low-income subgroup). Gaps for Asian students are discussed separately because this subgroup typically performs at higher levels than the other subgroups, including white students.

We carried out this study with advice from a panel of five nationally known experts in educational testing or education policy.<sup>1</sup> We also incorporated the following unique features into the design of the study:<sup>2</sup>

- **Data used in the study.** To conduct our achievement studies over the past three years, CEP has compiled state student achievement data from all 50 states with considerable technical support from our contractor for this project, the Human Resources Research Organization (HumRRO). Data were collected with the cooperation of state education officials, who verified the data's accuracy.
- **Availability of individual state profiles.** Using the extensive data collected for this study, CEP and HumRRO created a pair of comprehensive profiles for each of the 50 states. The first, a general trends profile, presents detailed information about achievement for students as a whole in a given state. The second, a subgroup achievement and gap trends profile, contains a wealth of data about the performance of subgroups, including data not discussed in this national report. Readers interested in particular states are encouraged to view and download these profiles from CEP's Web site ([www.cep-dc.org](http://www.cep-dc.org)).

<sup>1</sup> Members of the expert panel include Laura Hamilton, senior behavioral scientist, RAND Corporation; Eric Hanushek, senior fellow, Hoover Institution; Frederick Hess, director of education policy studies, American Enterprise Institute; Robert L. Linn, professor emeritus, University of Colorado; and W. James Popham, professor emeritus, University of California, Los Angeles.

<sup>2</sup> More details about study methods can be found in appendix 1 of part 1 of this series of reports, *Is the Emphasis on "Proficiency" Shortchanging Higher- and Lower-Achieving Students?* (CEP, 2009a), available at [www.cep-dc.org](http://www.cep-dc.org).

- **Use of trend lines.** State tests vary greatly in difficulty, content, scoring scales, cut scores for proficiency, and other aspects. Because there is no single, common metric for comparing results from these diverse tests, a different approach must be used to construct a national picture of achievement from the test results of many states. Our main unit of analysis is the trend line, which is a record of change in the performance of a specific subgroup at a particular achievement level in one subject and grade level for each year since 2002. For example, the change from 2002 to 2008 in the percentage proficient in reading for Latino 4<sup>th</sup> graders in Louisiana constitutes one trend line. Theoretically, the maximum number of possible trend lines for this study would be 5,400—two subjects, three grade levels, three achievement levels, and six student subgroups for each of the 50 states. In reality, however, fewer trend lines were available for analysis because certain subgroups were too small in some states to allow for a valid determination of trends, because comparable data were not available for a long enough period to constitute a trend, or because data were unavailable or missing for other reasons.
- **Years included in trends.** The subgroup trends analyzed for this report extend from 2002 (or later in some states) through 2008 and represent the most recent cycle of test results reported for NCLB by the time our data collection ended in April 2009. We used 2002 as the starting point because many states did not break out their test results by student subgroup and achievement level until they were required to do so by NCLB, and because trends in the limited number of states with comparable pre-2002 data were discussed in our 2007 report (CEP, 2007). This year, as in past years, only trend lines that encompassed at least three years of comparable test data for a particular subject, grade, and achievement level were included in our analyses. States with at least three years of comparable data but fewer than the full seven years were included as long as their data extended through 2007-08.
- **Use of average annual gains.** For the analysis by achievement level, we calculated an average annual percentage point gain or decline for each trend line. This is simply the increase or decrease in the percentage of students scoring at or above a certain achievement level, divided by the number of years of testing minus one (because we are looking at the difference between two years). In Arkansas, for example, 54% of low-income high school students scored at or above the proficient level in reading in 2008, compared with 29% in 2003—an overall gain of 25 percentage points. The average annual gain is therefore 5 percentage points per year—25 percentage points divided by five jumps between years of testing. To determine whether achievement gaps narrowed or widened, we compared the average annual percentage point gain made by a target subgroup with the average annual gain of the comparison group in the same state, subject, and grade level. If the average gain for the target subgroup was larger than that of the comparison group, we counted this as one instance of an achievement gap narrowing; if the average gain for the target group was smaller than that of the comparison group, we counted it as an instance of an achievement gap widening. This is a slightly different method than we used in previous reports, where we compared achievement gaps in the initial year of a trend line to the final year. The new method helped us capture more detailed information about whether gaps were narrowing because both the comparison and target groups were improving or because the comparison group showed a decline.
- **Use of mean scores.** In this study we used mean test scores, which are simply average test scores for groups of students, as an additional measure of achievement gaps. Mean scores have the advantage of reflecting achievement across the achievement spectrum, not just at the proficient level. The analysis of mean scores was similar to our analysis of per-

centages proficient in most key respects. To measure achievement gaps using mean scores, we subtracted the initial year mean score from the 2008 mean score for each subgroup then divided that figure by the number of intervening years. If the change in the mean score was greater for the target subgroup than for the comparison subgroup in the same state, subject, and grade level, we counted this as an instance of a gap narrowing.

## Grade 4 Subgroup Trends at Three Achievement Levels

Under NCLB, states must report student achievement at a minimum of three levels—basic, proficient, and advanced. (Students who fall below their state’s benchmark for basic achievement are considered “below basic,” a de facto fourth category.) NCLB gave states the latitude to define these levels in terms of their own tests and academic content standards; as a result, states’ definitions vary considerably.<sup>3</sup> Although the percentage of students scoring at or above the proficient level is the statistic used to determine adequate yearly progress under NCLB, the percentages scoring at the basic and advanced levels can also reveal some interesting trends about the performance of subgroups. For example, if the percentage of Latino students reaching or exceeding the basic level in math has increased over time, this indicates that more Latino students have improved enough to move from the below-basic into the basic category—progress that wouldn’t show up if one looked only at the percentage proficient. If the percentage of Native American students scoring at the advanced level in reading has fallen, this could signal that the highest-achieving students in this subgroup may have been neglected to some extent as educators work intensively to boost the percentage proficient. Or, if declines at the proficient and advanced levels were more prevalent for the African American subgroup than for the white subgroup, this would suggest that NCLB’s goal of raising achievement for all subgroups is not being met.

To better understand how various subgroups were performing across the achievement spectrum, we analyzed trends on state tests at three achievement levels for 4<sup>th</sup> graders from five racial-ethnic subgroups—African American, Asian, Latino, Native American, and white students—and for low-income and non-low-income students. In particular, we looked at the percentages of 4<sup>th</sup> graders in each subgroup scoring at the proficient level and above, the basic level and above, and the advanced level. Under this approach, the percentage proficient and above also includes students reaching the advanced level, and the percentage basic and above also includes students reaching the proficient and advanced levels. (Since there is no achievement level above advanced, the percentage advanced is a discrete category.) Using these cumulative achievement categories, rather than the discrete categories of basic alone or proficient alone, is consistent with how adequate yearly progress is determined under NCLB and is a simpler way to interpret trends that can become quite complex.

We limited this analysis to one elementary grade because of the massive amount of data involved and because this is the pilot year for a process that CEP hopes to extend to the middle and high school levels in future years. We looked at changes between school year 2001-02 (or the first year after that for which a state had comparable data) and school year 2007-08.

<sup>3</sup> In addition, some states use different names for the three achievement levels, and some states have established four or five such levels. More information about the process and challenges of analyzing performance by achievement level, as well as a detailed discussion of trends at three levels for students as a whole, can be found in part 1 of this series of CEP reports, *Is the Emphasis on “Proficiency” Shortchanging Higher- and Lower-Achieving Students?* (CEP, 2009a).

Unlike the gap trends discussed later in this report, trends by achievement level for a particular subgroup are not matched up with those from a comparison group (white or non-low-income students). We simply looked at the number of states in which particular subgroups have shown gains or declines in the percentages of students scoring at the basic-and-above, proficient-and-above, or advanced levels on state tests. To determine these trends, we used average annual percentage point gains, as explained in the previous section.

**TREND: Subgroups have made progress in grade 4 at all three achievement levels—basic-and-above, proficient-and-above, and advanced. In both reading and math, more states showed gains than declines for all subgroups analyzed at all three achievement levels.**

**Table 1** displays the numbers and percentages of states with gains, declines, or no net change for each subgroup at each achievement level in reading. **Table 2** presents the same information for math. For the Latino subgroup in reading, for example, 68% of the states with sufficient data showed gains in the percentage reaching the basic-and-above level, 84% made gains at the proficient-and-above level, and 62% had gains at the advanced level. More detailed tables showing the size of the gains (moderate-to-large or slight) for each of the three achievement levels can be found on the CEP Web site at [www.cep-dc.org](http://www.cep-dc.org). Based on advice from the expert panel, we classified gains or declines of at least one percentage point annually as moderate-to-large, and gains or declines of less than one percentage point annually as slight.

Achievement at the advanced level has received attention from researchers because of concerns about whether NCLB's focus on the proficient level has harmed higher-achieving students. In a previous report in this series (CEP, 2009a), we investigated this issue for students as a whole and found no strong evidence that higher-achieving students were being shortchanged; upward trends at the advanced level were much more common than downward trends.

For this current report, we analyzed trends at the advanced level at grade 4 by subgroup and found more states with gains than declines for all subgroups in both reading and math. We did, however, find differences by subject. As discussed below, a significant share of states with adequate data showed declines in the percentages of students reaching the advanced level in reading.

We also examined data from specific states to see which, if any, had concentrations of declines at the advanced level. For the most part, the declines were spread out rather than concentrated in a few states. Most states showed gains for most subgroups but had declines for one, two, or three subgroups in a particular subject. There were a few exceptions. Three states—Minnesota, Missouri and Ohio—had declines among 4<sup>th</sup> graders at the advanced level in at least 7 of the 12 possible trend lines (six subgroups times two subjects). In two states, Colorado and Oklahoma, the percentage of 4<sup>th</sup> graders reaching the advanced level increased for every subgroup in math but went down for every subgroup in reading. Without a detailed study of their student demographics, curriculum, instruction, and testing programs, it is inappropriate to draw conclusions about why these declines have occurred in certain states.<sup>4</sup>

<sup>4</sup> For more details about the states discussed here, see their state profiles of subgroup achievement and gap trends at [www.cep-dc.org](http://www.cep-dc.org).

**Table 1. READING: Number and percentage of states showing various trends for subgroups at three achievement levels on state grade 4 tests**

For years between 2002 and 2008 with comparable data

Achievement Level and Trend	African American	Asian	Latino	Native American	White	Low-Income
<b>Basic-and-above</b>						
Gain	18 (69%)	11 (48%)	19 (68%)	8 (50%)	18 (53%)	24 (71%)
Decline	7 (27%)	9 (39%)	8 (29%)	6 (38%)	9 (26%)	9 (26%)
No change	1 (4%)	3 (13%)	1 (4%)	2 (13%)	7 (21%)	1 (3%)
Total # of states with data	26	23	28	16	34	34
<b>Proficient-and-above</b>						
Gain	31 (86%)	26 (79%)	32 (84%)	16 (80%)	37 (84%)	33 (77%)
Decline	5 (14%)	6 (18%)	6 (16%)	4 (20%)	5 (11%)	7 (16%)
No change	0 (0%)	1 (3%)	0 (0%)	0 (0%)	2 (5%)	3 (7%)
Total # of states with data	36	33	38	20	44	43
<b>Advanced</b>						
Gain	21 (60%)	22 (69%)	23 (62%)	12 (63%)	28 (67%)	26 (63%)
Decline	11 (31%)	9 (28%)	12 (32%)	5 (26%)	11 (26%)	11 (27%)
No change	3 (9%)	1 (3%)	2 (5%)	2 (11%)	3 (7%)	4 (10%)
Total # of states with data	35	32	37	19	42	41

Table reads: Of the 26 states with sufficient data to analyze trends for African Americans, 18 states, or 69%, showed gains in the percentage of African American students scoring at or above the basic level of achievement in reading. Seven of these states, or 27%, showed declines for African Americans at the basic-and-above level in reading, and one state, or 4%, showed no net change.

NOTE: Percentages in some columns do not total 100% due to rounding.

**Table 2. MATH: Number and percentage of states showing various trends for subgroups at three achievement levels on state grade 4 tests**

For years between 2002 and 2008 with comparable data

Achievement Level and Trend	African American	Asian	Latino	Native American	White	Low-Income
<b>Basic-and-above</b>						
Gain	22 (81%)	15 (63%)	25 (86%)	12 (71%)	27 (77%)	28 (80%)
Decline	4 (15%)	4 (17%)	4 (14%)	4 (24%)	4 (11%)	7 (20%)
No change	1 (4%)	5 (21%)	0 (0%)	1 (6%)	4 (11%)	0 (0%)
Total # of states with data	27	24	29	17	35	35
<b>Proficient-and-above</b>						
Gain	32 (89%)	29 (88%)	36 (95%)	15 (71%)	39 (89%)	37 (84%)
Decline	4 (11%)	4 (12%)	2 (5%)	4 (19%)	5 (11%)	6 (14%)
No change	0 (0%)	0 (0%)	0 (0%)	2 (10%)	0 (0%)	1 (2%)
Total # of states with data	36	33	38	21	44	44
<b>Advanced</b>						
Gain	32 (91%)	29 (91%)	30 (81%)	16 (80%)	36 (86%)	33 (79%)
Decline	1 (3%)	3 (9%)	5 (14%)	4 (20%)	4 (10%)	6 (14%)
No change	2 (6%)	0 (0%)	2 (5%)	0 (0%)	2 (5%)	3 (7%)
Total # of states with data	35	32	37	20	42	42

Table reads: Of the 27 states with sufficient data to analyze trends for African Americans, 22 states, or 81%, showed gains in the percentage of African American students scoring at or above the basic level of achievement in mathematics. Four of these states, or 15%, showed declines for African Americans at the basic-and-above level in math, and one state, or 4%, showed no net change.

NOTE: Percentages in some columns do not total 100% due to rounding.

**TREND: Subgroup results were more positive in math than in reading at all achievement levels and were especially notable in math at the advanced level for African American students.**

Although positive trends outweighed negative trends for every subgroup at all three achievement levels, results did vary by subject and subgroup.

In math, very high percentages of the states with sufficient data—80% or more of these states—posted gains at grade 4 for nearly all subgroups at both the proficient and advanced levels. (See table 2.) Of particular note are the high numbers of states that showed math gains for the Latino subgroup at the proficient level (95% of the states with sufficient data) and for the African American subgroup at the advanced level (91%). In both of these cases, the percentage of states that showed progress for Latino or African American students was equal to or exceeded the comparable percentage for the Asian subgroup and was higher than that for the white subgroup.

In reading, the proportions of states with gains were highest at the proficient-and-above level. (See table 1.) A majority of the states with data also showed gains in reading, mostly moderate-to-large in size, at the advanced level for all subgroups. Still, a sizeable share of the states with data, ranging from 26% of these states for Native American and white students to 32% for Latino students, experienced declines in the percentages reaching advanced. (The majority of these declines were slight, however.)

In reading, some variations by subgroup were also evident. At the basic-and-above level, gains in reading were more prevalent for African American, Latino, and low-income students than for the other subgroups. For example, 71% of the states with sufficient data had increases in the percentage of low-income students at the basic-and-above level, while 48% of states had increases at this achievement level for Asian students. At the advanced level, slightly higher proportions of states had gains in reading for the Asian and white subgroups than for the other subgroups. At the proficient level, the percentages of states with gains did not vary greatly among subgroups or fit into a clear pattern.

The differences we noted in reading may stem partly from the fact that we used percentages of states to compare a relatively small number of cases. At the basic-and-above level, for example, 23 states had data for the Asian subgroup and just 16 states had data for the Native American subgroup. If only two more states had demonstrated increases rather than declines for Native Americans, the percentage of states with gains would have risen from 50% to 63%. Comparing percentages of states showing various trends is a blunt instrument but a necessary one, given the lack of a common metric for aggregating results from different state tests.

## Gap Trends Based on Percentages Proficient at Three Grade Levels

We analyzed gaps in the percentages of students scoring proficient or above on state tests at three grade levels—grade 4, middle school (almost always grade 8),<sup>5</sup> and the high school grade tested for NCLB accountability, which varies by state. We looked specifically at gaps between African American, Latino, or Native American students and white students, and between low-

<sup>5</sup> Grade 8 was used in all states except Utah, which administers an end-of-course exam in algebra at various grade levels as its middle school math test.

income and non-low-income students. Our trend lines extended from 2002 (or the first year after that for which a state had comparable data) through 2008. We did a separate analysis of achievement for the Asian subgroup for reasons explained at the end of this section.

**TREND: Achievement gaps have narrowed in most states at the elementary, middle, and high school levels, although in a sizeable minority of cases gaps have widened. Across all subgroups, grade levels, and subjects analyzed, 74% of the trend lines showed gaps narrowing, and 23% showed gaps widening. In the remainder, the gaps showed no change.**

**Table 3** displays the results for grade 4, the elementary grade analyzed for this study. In reading, for example, the grade 4 achievement gap between African American and white students narrowed in 28 of the 36 states with sufficient data, widened in 7 states, and showed no net change in 1 state. The Latino-white gap narrowed in 29 of the 38 states with data and widened in 9 states. The gap between low-income and non-low-income students shrunk in 31 of 43 states and grew in 12 states.

<b>Table 3. GRADE 4: Number of states showing various trends in percentage proficient gaps on state reading and math tests</b>					
<b>Trend</b>	<b>African American &amp; White</b>	<b>Latino &amp; White</b>	<b>Native American &amp; White</b>	<b>Low-Income &amp; Not Low-Income</b>	<b>Total Trend Lines</b>
<b>Reading</b>					
States in which gap narrowed	28	29	12	31	100
States with no change in gap	1	0	0	0	1
States in which gap widened	7	9	7	12	35
States with sufficient trend data	36	38	19	43	136
<b>Math</b>					
States in which gap narrowed	29	29	11	32	101
States with no change in gap	2	0	1	1	4
States in which gap widened	5	9	9	11	34
States with sufficient trend data	36	38	21	44	139

Table reads: Of the 36 states with sufficient data to analyze trends for African Americans, the gap between African American and white students in the percentage scoring proficient in grade 4 reading narrowed in 28 states, showed no change in 1 state, and widened in 7 states. Across all subgroups, 100 of the 136 trend lines analyzed in grade 4 reading showed gaps narrowing.

For the African American, Latino, and low-income subgroups, more states showed gaps narrowing than widening at grade 4 in both reading and math. For the Native American-white gap, the counts were closer, although this gap still narrowed in more states.

At the middle school level, percentage proficient gaps also narrowed far more often than they widened, as shown in **table 4**. For instance, the Latino-white gap narrowed in reading in 28 of the 36 states with sufficient data.

High school results (**table 5**) are less positive at first blush, but this is complicated by the larger number of states that lacked sufficient data. One reason for the smaller pool of data at the high school level is that some states use end-of-course exams. For instance, they may administer multiple math exams in algebra I, algebra II, and geometry that individual students take at different grades after they complete the appropriate course. When we asked state officials to select the most appropriate test to capture high school performance trends, some were unwilling or unable to do so.

Our overall finding that achievement gaps have narrowed in most states is consistent with a general finding from other studies that gaps have been slowly shrinking over the past few decades on the National Assessment of Educational Progress and other standardized measures. These long-term trends are described in **box A**.

**Table 4. MIDDLE SCHOOL: Number of states showing various trends in percentage proficient gaps on state reading and math tests**

Trend	African American & White	Latino & White	Native American & White	Low-Income & Not Low-Income	Total Trend Lines
<b>Reading</b>					
States in which gap narrowed	30	28	14	30	102
States with no change in gap	0	2	2	4	8
States in which gap widened	5	6	5	8	24
States with sufficient trend data	35	36	21	42	134
<b>Math</b>					
States in which gap narrowed	26	27	14	30	97
States with no change in gap	2	0	2	3	7
States in which gap widened	7	8	6	10	31
States with sufficient trend data	35	35	22	43	135

Table reads: Of the 35 states with sufficient data to analyze trends for African Americans, the gap between African American and white students in the percentage scoring proficient in middle school reading narrowed in 30 states and widened in 5 states. Across all subgroups, 102 of the 134 trend lines analyzed in middle school reading showed gaps narrowing.

**Table 5. HIGH SCHOOL: Number of states showing various trends in percentage proficient gaps on state reading and math tests**

Trend	African American & White	Latino & White	Native American & White	Low-Income & Not Low-Income	Total Trend Lines
<b>Reading</b>					
States in which gap narrowed	19	24	13	25	81
States with no change in gap	3	0	1	0	4
States in which gap widened	6	4	3	10	23
States with sufficient trend data	28	28	17	35	108
<b>Math</b>					
States in which gap narrowed	21	23	11	21	76
States with no change in gap	1	1	1	2	5
States in which gap widened	6	4	4	10	24
States with sufficient trend data	28	28	16	33	105

Table reads: Of the 28 states with sufficient data to analyze trends for African Americans, the gap between African American and white students in the percentage scoring proficient in high school reading narrowed in 19 states, showed no net change in 3 states, and widened in 6 states. Across all subgroups, 81 of the 108 trend lines analyzed in high school reading showed gaps narrowing.

**Box A. Long-term trends in achievement gaps**

The general trend we observed of achievement gaps on state tests decreasing since 2002 can be seen as part of a longer trend going back well before NCLB was enacted. According to nationwide data from the National Assessment of Educational Progress and other measures, achievement gaps have been shrinking for several decades.

Gaps have been narrowing on NAEP since the long-term NAEP tests were first administered in the 1970s, as shown in **table 6**. The gaps in NAEP scores between African American and white students at ages 9, 13, and 17 have shrunk overall since the 1970s but with some waxing and waning over the years (NCES, 2009). The black-white gap was at its smallest in the late 1980s or early 1990s, depending on the subject and student age group tested, and had increased somewhat by 1999 (Harris & Herrington, 2006). Hedges & Nowell (1998) also found that gaps on NAEP narrowed overall between 1971 and 1994; the black-white gap was narrowest in 1988 but increased slightly thereafter. Even with the general narrowing trend, the black-white and Latino-white gaps on NAEP remain large—anywhere from 16 to 29 points on the NAEP scoring scale of 0-500, as illustrated in table 6.

According to Grissmer, Flanagan, and Williamson (1998), African American-white gaps have narrowed because African American scores rose immensely over a 25-year time period while scores for white students increased only slightly. For example, over roughly 25 years, NAEP scores for African American high school students improved by about 0.6 standard deviations in math and reading, while NAEP scores for white high school students improved by roughly 0.1 standard deviations. Again, the late 1980s were the height of improvement. Gaps were found to have narrowed at the elementary and middle school levels as well.

*Continued on next page*

**Table 6. Size of achievement gaps in scale scores\* on the long term NAEP, 1970s<sup>†</sup> through 2008**

Subgroup, Subject, Age	Initial Score*	Gap with White Students	2008 Score	Gap with White Students
African American, math, age 9	192	32	224	26
African American, math, age 13	230	42	262	28
African American, math, age 17	268	38	287	27
African American, reading, age 9	181	36	204	24
African American, reading, age 13	226	36	247	21
African American, reading, age 17	241	52	266	29
Latino, math, age 9	203	21	234	16
Latino, math, age 13	238	34	268	22
Latino, math, age 17	276	30	293	21
Latino, reading, age 9	183	34	207	21
Latino, reading, age 13	232	30	242	26
Latino, reading, age 17	252	41	269	26

\*The long-term NAEP is scored on a scale of 0-500.

<sup>†</sup>The initial year of the trend line varies by subgroup, between 1971 and 1978.

Source: U.S. Department of Education, *National Assessment of Educational Progress, Long-Term Trend Assessments, 1978, 1982, 1986, 1990, 1992, 1994, 1996, 1999, 2004, and 2008*, <http://nces.ed.gov/nationsreportcard/ltt/>.

In recent decades, achievement gaps between African American and white students have narrowed on other measures as well, including General Social Survey vocabulary tests and IQ tests (Huang & Hauser, 2000; Dickens & Flynn, 2006). Hedges & Nowell (1998) found black-white test scores converging on six different measures other than NAEP, although the convergence in scores tended to take place more at the lower end of the distribution than at the top—in other words, the gap between lower-performing black and white students was shrinking but tended to remain large among the highest-achieving black and white students. Several other indicators also show African American-white achievement gaps narrowing over the last few generations (Jencks & Phillips, 1998).

As shown in table 6, Latino-white achievement gaps have also been narrowing over time, although these gaps remain large as well. Some studies have concluded that the Latino-white gap tends to shrink as students progress through school (Fryer & Levitt, 2004; Clotfelter, Ladd, & Vigdor, 2006), while others have found that the gap remains constant (National Research Council, 2006). However, the generic terms “Latino” and “Hispanic” may mask considerable differences within this diverse subgroup. For example, U.S. poverty rates are far lower for students of Cuban or South American origin than for students of Dominican, Puerto Rican, or Mexican origin (Reardon & Galindo, 2008). On the other hand, students of Puerto Rican origin tend to be far more proficient in English than Dominicans. And, not surprisingly, third-generation Latinos are generally far more likely to be wealthier and more proficient in English than recent arrivals. According to Reardon and Galindo (2008), the largest achievement gaps with white students are found among lower-income first- and second-generation students from Mexico and Central American nations who are not proficient in English and who speak only Spanish at home.

Media reports have recently highlighted a widening of gaps between groups of students on the SAT college admissions test (Hechinger, 2009)—a trend that runs counter to most long-term studies of achievement gaps. While the SAT may provide useful information about the academic skills of college-bound students in a given year, it is not the most appropriate tool for understanding achievement gaps in K-12. First, students who take SATs are not a sample of all high school students, so the College Board recommends

*Continued on next page*

against using SAT scores to compare achievement for states, schools, districts, or subgroups (College Board, 2009a). Second, the pool of students taking the SAT has become larger and far more diverse as more students aspire to attend college. Many states have programs that encourage or even require high school students to take the SAT, and even help pay test fees (Education Commission of the States, 2008). The College Board has noted an increase, for example, in the percentage of Latino students and students whose first language is not English who take the SAT (College Board, 2009b). As a result of this expansion, the pool of SAT test-takers now includes a greater variety of students at different achievement levels than in the past. This could account for some of the widening achievement gap on the SAT.

Despite the prevalence of narrowing gaps, it is troubling that 23% of the state trend lines we studied showed achievement gaps widening. For example, the percentage proficient gap between low-income and non-low-income students in high school math increased in 10 of the 33 states with sufficient data. The Native American-white gap in elementary school math widened in 9 of 21 states.

We searched the data for any patterns among states with widening gaps. For the most part, instances of gaps widening were scattered throughout the states rather than concentrated in certain states. Usually, most of gaps in a state narrowed except for one or two instances. For example, in Wisconsin, the percentage proficient gap between Native American and white students decreased in all grade level/subject combinations except grade 4 reading.

We did identify some states in which gaps widened for the same subgroup in at least four out of six possible trend lines (two subjects times three grade levels). In Missouri, the African American-white gap increased in four out of six instances. In Minnesota, the Latino-white gap grew across the board. Alaska, Texas, and Washington had four or more instances of gaps widening for the Native American subgroup. Six states—Maine, Minnesota, Missouri, South Dakota, Washington, and Wisconsin—showed four or more instances of gaps widening for the low-income subgroup.

One must refrain, however, from drawing conclusions about educational effectiveness in states with several examples of widening gaps. First, achievement gap trends are affected by a variety of demographic factors, such as the number of students who are recent arrivals to the U.S. and are English language learners. More in-depth research is needed to explain why achievement gaps are widening in certain states. Second, in many of these cases, the achievement of target subgroups is still improving, although not as quickly as that of the white or non-low-income subgroups—hence the widening gaps. In Missouri, for example, African American students made gains in middle school reading and math, as well as in high school math, but these gains were slightly smaller than those of white students.

**TREND: Gaps have typically narrowed because the achievement of lower-performing subgroups has gone up rather than because the achievement of higher-performing subgroups has gone down.**

Gaps can narrow not only because achievement has gone up for racial/ethnic minority students but also because white student achievement has gone down, or because both minority and white students have lost ground but the white subgroup has declined at a greater rate. The same dynamics can occur between low-income and non-low-income students. Have gaps narrowed in many cases because the performance of white students or non-low-income students has declined?

We investigated this question and found that generally, this was not the case. In this study, gaps most often narrowed because gains made by target subgroups outpaced gains by white or non-low-income students—in other words, both subgroups improved, but the target subgroup improved more than the comparison group. Conversely, we found that gaps most often widened because both groups improved, but the comparison subgroup improved more than the target group.

**Table 7** displays the results of our investigation. For example, across all grade levels and subjects in states with sufficient data, we identified a total of 153 trend lines that showed African American-white gaps narrowing. In 142 of these instances, the gap narrowed because, although both groups made gains, the percentage proficient grew at a faster rate for African American students than for white students. In only seven instances did the gap narrow because the white percentage proficient declined while the African Americans percentage proficient increased. Finally, in four instances, both African American and white student performance declined, but the drop was greater for white students.

An example from this same subgroup shows the patterns that can lead to widening gaps. In 36 instances across all grade levels and subjects, the African American-white gap widened. In 19 of these instances, gaps increased because the performance of both subgroups improved, but the white subgroup made larger gains than the African American subgroup. In 10 instances, the white subgroup made gains, but the African American subgroup declined. And in seven instances, both subgroups showed declines, but the decrease was greater for African Americans.

**Table 7. Details on narrowing and widening gaps in percentages proficient**

Gap Trend	African American & White	Latino & White	Native American & White	Low-Income & Not Low-Income
Total instances of gaps narrowing	153	160	75	169
Both groups improved, target group improved more	142	149	63	155
Comparison group declined, target group improved	7	5	8	7
Both groups declined, comparison group declined more	4	6	4	7
Total instances of gaps widening	36	40	34	61
Both groups improved, comparison group improved more	19	25	20	39
Comparison group improved, target group declined	10	8	7	9
Both groups declined, target group declined more	7	7	7	13

Table reads: A total of 153 trend lines were analyzed that showed African American-white gaps narrowing. In 142 instances, gaps narrowed because both subgroups improved but the African American subgroup improved more. In seven instances, the white percentage proficient declined but the African American percentage improved. In four instances, the percentage proficient fell for both subgroups but white student achievement decreased more.

### TREND: Gaps narrowed more often for the Latino and African American subgroups than for other subgroups.

To make it easier to compare the progress of various subgroups in closing achievement gaps, we pulled out the instances of gaps narrowing across all states from tables 3, 4, and 5, and then broke them down by subgroup and grade level. **Table 8** presents the results. For example, among all the states with sufficient data to compare gaps between low-income and non-low-income students, there were a total of 240 gap trend lines for low-income students across all grade levels and subjects. In 70% of these instances, the gap between low-income and non-low-income students narrowed.

For nearly all subjects and grade levels, gaps narrowed more often for the African American and Latino subgroups than for the Native American or low-income subgroups. In math, for example, the African American-white gap narrowed in 77% of the instances we studied, and the Latino-white gap narrowed 78% of the time. This compares with 61% of instances for the Native American subgroup and 69% for the low-income subgroup. Across both subjects and three grade levels, gaps narrowed 77% of the time for African Americans and 79% for Latinos, compared with 65% for Native Americans and 70% for low-income students.

For African Americans and low-income students, gaps narrowed more often at the elementary and middle school levels than at the high school level. While 79% of the gaps between African-American and white students narrowed at grade 4 and 80% did so at the middle school level, this figure went down to 71% at the high school level. By contrast, the percentage of gaps narrowing went up between the elementary and high school levels for Latinos and Native Americans.

Two caveats are in order about making comparisons across subgroups. First, there is some overlap among these subgroups due to NCLB requirements. For example, if a Latino student qualifies for free or reduced-price lunch, that student must be counted in both the Latino and low-income subgroups. Second, the pool of states with sufficient data to make comparisons between the Native American and white subgroups is far smaller than for the

**Table 8. Percentage of gaps narrowing by subgroup, grade level, and subject**

Trend	African American & White	Latino & White	Native American & White	Low-Income & Not Low-Income
% narrowing at grade 4	79%	76%	58%	72%
% narrowing at middle school	80%	77%	65%	71%
% narrowing at high school	71%	84%	73%	68%
% narrowing in reading across all grade levels	78%	79%	68%	72%
% narrowing in math across all grade levels	77%	78%	61%	69%
Total % narrowing by subgroup	77%	79%	65%	70%
Total instances	198	203	116	240

Table reads: At grade 4, the gaps in percentages proficient between African American and white students narrowed in 79% of the instances analyzed, while the gaps between Latino and white students narrowed in 76% of instances.

other subgroups, so there were notably fewer instances of Native American-white gaps for our analysis. This is because the Native American population is concentrated in a limited number of states, while the other subgroups are spread more evenly throughout the country. If just a small number of gap trends were different (say, if a gap had narrowed instead of widened), this would have a greater impact on the total percentage of gaps for Native Americans than for other subgroups.

### **TREND: Gaps are still large.**

Despite the progress being made, the nation still has a long way to go to close achievement gaps between student subgroups. To get a rough idea of the size of the gaps, we averaged the percentage proficient for each subgroup across all states with data and examined the differences between groups. It is important to recognize that this is not a nationwide percentage proficient figure for either group; rather, it is the unweighted average across the 28 states with data to make this comparison.

Generally, the gaps were largest in high school math and smallest in elementary school math. Overall, the average gaps in percentages proficient across states were largest for the African American subgroup. In high school math, for example, the mean (average) percentage proficient was 45% for the African American subgroup and 74% for the white subgroup, resulting in a black-white gap of 29 percentage points. The high school math gaps were 22 to 23 percentage points for Latinos, Native Americans, and low-income students. The Asian subgroup outperformed the white subgroup in high school math by three percentage points.

These types of averages cover up the huge differences between states in percentages proficient for each subgroup. These differences are probably explained largely by differences in the difficulty of state tests and the location of cut scores for proficient performance—differences that reflect the fact that almost all of the 50 states have created their own standards and testing programs, as permitted under NCLB. What constitutes “proficiency” in one state may vary considerably from the definition in another state. In Minnesota, 11% of students in the Native American subgroup are considered proficient in high school math, while in Alabama, 90% of the Native American subgroup is proficient.

These variations in state standards and tests can lead to very different conclusions about the size of achievement gaps. For example, 91% of Latinos in Georgia were proficient in high school math, compared with 96% of white students—a difference of five percentage points. In neighboring South Carolina, 50% of high school Latinos were proficient in math, compared with 73% of their white peers—a 23 percentage point difference. These differences in the number and percentage of students at the proficient level (and the basic and advanced levels as well) occur partly because gaps may appear narrower or wider depending on where a state has set a cut score on the scoring scale for its test. Simply put, if a cut score on a test is very high or very low (so almost everybody or almost nobody reaches it) there is little apparent gap in percentage proficient terms. A cut score closer to the mean test score will be more sensitive to detecting gaps in percentages proficient, and the gaps between subgroups will appear larger. This phenomenon is explored in more detail in the section on gaps in mean scores.

**TREND: In the vast majority of states with sufficient data, the Asian subgroup outperformed all other subgroups, including white students, in all subject and grade level combinations except high school reading.**

We examined gaps between the Asian and white subgroups separately because the Asian subgroup tends to outperform the white subgroup academically. These two subgroups were typically the highest-scoring of the groups we analyzed. In many cases, however, the percentage of Asian students reaching the proficient level started out higher and then grew at a faster rate than that of the white subgroup. Taken literally, this means that the white-Asian gap has widened. But if we had counted these situations as instances of gaps widening, it would have distorted the national picture of progress in closing achievement gaps.

Still, it is important to understand how the Asian subgroup is doing on state tests. **Table 9** shows the number of states in which the Asian subgroup had a higher percentage proficient than the white subgroup, as well as the total number of states with adequate data to make such a comparison. In grade 4 math, for example, the percentage proficient for Asian students exceeded that for white students in 28 of the 34 states with adequate data. This pattern of Asians outperforming whites was consistent for every subject and grade level combination except high school reading, where Asians outperformed whites in just 7 of 27 states.

The Asian subgroup was the only racial/ethnic subgroup that typically outperformed white students. The other subgroups scored higher than the white subgroup in only two instances: in Nebraska, the Latino subgroup had a higher percentage proficient in 2008 than the white subgroup in middle school math; and in Oklahoma, Native American students had a higher percentage proficient than white students in grade 4 reading.

While this initial picture of Asian subgroup achievement is very positive, it should be noted that the percentages proficient for this subgroup vary considerably from state to state. In Pennsylvania, for example, 78% of the Asian subgroup reached proficiency in high school math in 2008, compared with 62% of the white subgroup, for a white-Asian achievement gap of 16 percentage points. But in Minnesota, the situation was reversed: only 52% of the Asian subgroup was proficient in math at the middle school level, compared with 63% of the white subgroup, for an Asian-white gap of 11 percentage points.

**Table 9. Number of states in which the Asian subgroup had a higher percentage proficient than the white subgroup, 2008**

Subject & Grade Level	States in Which Asians Outperformed Whites	States with Adequate Data
Grade 4 reading	21	34
Middle school reading	20	32
High school reading	7	27
Grade 4 math	28	34
Middle school math	29	32
High school math	23	28

Table reads: In 21 of the 34 states with sufficient data for this analysis, a higher percentage of Asian students than of white students scored at or above the proficient level in grade 4 reading in 2008.

These types of discrepancies may be partly due to differences among states in the composition of the Asian subgroup. The catch-all term “Asian” may include students whose families come from nations as diverse as Pakistan and Japan and may lump together very recent immigrants with students whose families have been in the U.S. for many generations. The educational needs of Hmong students in Minnesota, Korean students in California, or native Hawaiian students may be entirely different. For example, 52% of Hmong students in the U.S. are English language learners, compared with slightly less than 8% of Filipino students. The “model minority” label that is sometimes applied to Asians covers up the fact that key segments of the Asian subgroup are struggling academically and face many of the same problems as do other groups of English language learners (Asian American Legal Defense and Education Fund, 2008; Li & Wang, 2008; and Zhao & Qiu, 2009).

Several reasons have been suggested for the superior academic achievement of Asians (or at least, specific ethnic groups within this broad category). These include deference to parental and teacher authority, less household television watching, a greater cultural emphasis placed on academic achievement, and higher parental involvement and educational and income levels. For reviews and critiques of this research, see Li & Wang, 2008; Zhao & Qiu, 2009; and Thernstrom & Thernstrom, 2004.

## Gap Trends Based on Mean Scores at Three Grade Levels

The No Child Left Behind Act requires states to use the percentage proficient indicator to determine whether districts and schools have made adequate yearly progress. This indicator has limitations, however, that are especially pertinent to analyses of achievement gaps. As noted above, gaps expressed as differences in percentages proficient may appear larger or smaller depending on where the proficient cut score for a test has been set—a phenomenon described in more detail below. Another limitation is that the percentage proficient does not provide information about how well students as a whole or in subgroups are performing at points on the achievement spectrum above or below proficient. For example, if a group of students has made rapid advances above the proficient level, this would not be captured by the percentage proficient indicator.

To address these problems and provide a check on the percentage proficient indicator, we also analyzed achievement gaps using a second indicator, mean scores. These are simply average test scores for groups of students expressed on the scoring scale for that particular test. For example, Nevada’s Criterion Referenced Test for grades 3-8 is scored on a scale of 100-500; the mean score for all 8th grade students who took that test in 2008 was 298.6. Mean scores avoid the cut score problem described above and capture achievement across the spectrum of performance, not just at the proficient level.

To measure achievement gaps using mean scores, we first subtracted the mean score for the starting year of the trend (2002 or later) from the mean score for 2008 for each subgroup. The difference represented the change in mean scores. We calculated an average annual gain and then compared the average changes in mean score for the target subgroup and the comparison subgroup. If the average change in mean score was greater for the target group than for the comparison group in the same state, subject, and grade level, this meant the achievement gap had narrowed. If the average change was greater for the comparison group than for the target group, this meant the gap had widened. The Asian subgroup was not included in these calculations because this group typically outperformed white students.

One problem with using mean scores was that they were somewhat more difficult to obtain from states, especially broken out by student subgroup. States are not required by NCLB to report test results in terms of mean scores, and many states do not make them available. For example, 36 states had sufficient years of comparable percentage proficient data to calculate the Latino-white gap in grade 4 reading, but just 31 states provided us with the mean scores for this same comparison. Consequently, our mean score analyses included fewer states.

**TREND: Mean scores for subgroups have risen in most cases.**

As we demonstrate below, mean scores do not show as much progress in closing achievement gaps as the percentage proficient indicator. But that does not mean test scores for student subgroups are declining. In fact, for all six subgroups discussed in this report, the majority of trend lines based on mean scores show improvement. For the African American-white gap comparisons, for example, we had sufficient mean score data to analyze 165 trend lines across all grade levels and both subjects. Just 22 of these trend lines, or 13%, showed declines in mean scores for African Americans. Mean scores declined in 13% of the trend lines analyzed for the Latino subgroup and 15% for the low-income subgroup. The percentage of trend lines showing declines in mean scores was higher for the Native American (24%) and Asian (24%) subgroups. In most cases where gaps have widened, it was because both the target and comparison groups made gains in mean scores, but the comparison group (white or non-low-income students) made greater gains. So the mean score gap trends described below are occurring in a context of overall improvement on state tests for all subgroups.

**TREND: Although mean scores show more gaps narrowing than widening, mean scores give a less rosy picture of progress in closing achievement gaps than percentages proficient.**

**Table 10** compares trends in achievement gaps by subgroup using the mean score indicator and the percentage proficient indicator for the subset of states with both types of data. Overall, gap trend lines for this subset of states narrowed 71% of the time using percentages proficient and 58% of the time using mean scores. And in 37% of the instances we looked at in these states, gaps widened according to mean scores—considerably more than the 24% of gaps that widened according to percentages proficient and a cause for concern.

Roughly one-third of the gaps in mean scores widened for the African American and Latino subgroups, and for low-income students this percentage was even higher at 38%. The starkest difference between the two indicators can be seen in gaps for the Native American subgroup. Using mean scores, we found more instances of Native American-white achievement gaps widening (51%) than narrowing (43%). However, many states did not have the data needed to make comparisons for this subgroup, so the number of trend lines analyzed was low—just 81.

We also broke out the data on mean score and percentage proficient gap trends by subject and grade level, as shown in **table 11**, and found some variations by grade level. In general, there was more divergence between the mean score and percentage proficient trends at the elementary level than at the middle and high school levels. In high school math, both indicators showed similar percentages of narrowing and widening gaps.

**Table 10. GAPS BY SUBGROUP: Comparison of percentage proficient and mean score measures of achievement gaps**

Target and Comparison Subgroups	Percentage Proficient (# and % of instances showing various trends)			Mean Scores (# and % of instances showing various trends)			Number of Trend Lines
	Narrowed	Widened	No Change	Narrowed	Widened	No Change	
African American & white	120 (75%)	31 (19%)	8 (5%)	104 (65%)	50 (31%)	5 (3%)	159
Latino & white	120 (75%)	38 (24%)	3 (2%)	99 (61%)	53 (33%)	9 (6%)	161
Native American & white	47 (58%)	27 (33%)	7 (9%)	35 (43%)	41 (51%)	5 (6%)	81
Low-income & not low-income	124 (70%)	45 (25%)	9 (5%)	96 (54%)	68 (38%)	14 (8%)	178
Total	411 (71%)	141 (24%)	27 (5%)	334 (58%)	212 (37%)	33 (6%)	579

Table reads: Gaps between African American and white students narrowed in 120 out of 159 instances, or 75% of the time, using the percentage proficient measure; this same gap narrowed in 104 instances, or 65% of the time, using the mean score measure.

**Table 11. GAPS BY SUBJECT AND GRADE LEVEL: Comparison of percentage proficient and mean score measures of achievement gaps**

Target and Comparison Subgroups	Percentage Proficient (# and % of instances showing various trends)			Mean Scores (# and % of instances showing various trends)			Number of Trend Lines
	Narrowed	Widened	No Change	Narrowed	Widened	No Change	
Grade 4 reading	74 (70%)	31 (29%)	1 (1%)	55 (52%)	43 (41%)	8 (8%)	106
Grade 4 math	80 (72%)	28 (25%)	3 (3%)	56 (50%)	45 (41%)	10 (9%)	111
Middle reading	74 (72%)	21 (20%)	8 (8%)	65 (63%)	33 (32%)	5 (5%)	103
Middle math	77 (72%)	23 (21%)	7 (7%)	63 (59%)	41 (38%)	3 (3%)	107
High school reading	54 (72%)	17 (23%)	4 (5%)	44 (59%)	27 (36%)	4 (5%)	75
High school math	52 (68%)	21 (27%)	4 (5%)	51 (66%)	23 (30%)	3 (4%)	77
Total	411 (71%)	141 (24%)	27 (5%)	334 (58%)	212 (37%)	33 (6%)	579

Table reads: Gaps in grade 4 reading narrowed in 74 out of 106 instances, or 70% of the time, using the percentage proficient measure; this same gap narrowed in 55 instances, or 52% of the time, using the mean score measure.

To better understand the differences between the two indicators, we examined the 217 trend lines available for grade 4 in more detail. We found that the majority of the time, in 156 instances, the mean score and percentage proficient indicators converged—that is, if gaps narrowed on one measure, they also narrowed on the other. In 62 instances, or more than one fourth of the time, the two indicators showed conflicting trends. In 33 of the 62 instances, the indicators conflicted because gaps narrowed according to the percentage proficient but widened according to the mean score. In 9 instances the reverse was true—gaps widened according to the percentage proficient but narrowed according to the mean score. In 19 instances, one indicator showed no change in the size of a gap and the other showed the gap widening or narrowing.

**TREND: The differences in gap trends that emerge when mean scores are used rather than percentages proficient highlight the impact of cut scores for proficiency on the size of achievement gaps.**

As noted above, analyses of mean scores pick up changes in achievement at the lower and higher ends of the scoring scale that are not captured by the percentage proficient indicator. For example, if one subgroup made larger gains than another at the advanced achievement level, the mean score for the first group could increase more than the mean for the second, thereby narrowing or widening the mean score gap but having no impact on the percentage proficient gap. The same thing could happen if one subgroup made greater gains than another in moving students from the below basic to the basic level (but not quite to the proficient level). As described in the first section of this report, subgroups have indeed made gains at the basic and advanced levels, and differences among subgroups in the relative size of these gains are undoubtedly affecting gaps in mean scores without influencing percentage proficient gaps.

But a statistical phenomenon relating to cut scores also comes into play when gaps are measured using the percentage proficient, and this phenomenon is crucial in explaining why gap trends look somewhat less positive when mean scores are used. As already noted, percentage proficient gaps may appear larger or smaller depending on where the cut score for proficiency has been set. With cut scores at the very high or low end of the scoring scale, gaps between subgroups will appear to be quite small; with cut scores closer to the mean for a test, gaps will appear larger. **Box B** illustrates how this phenomenon works.

## Conclusion

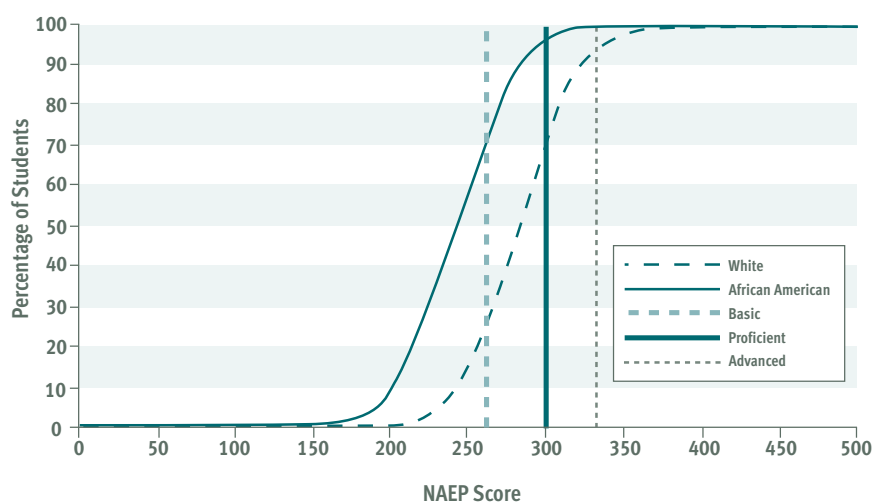
The good news from this study is that overall, states have made progress in closing achievement gaps, as measured by the percentage of students in various subgroups reaching the proficient level on state tests. In 74% of all cases, gaps in percentages proficient narrowed. In other words, the improvements made by target subgroups were larger than the gains shown by comparison subgroups. Also, many people tend to assume the worst when they hear that achievement gaps are widening, as was the case in 23% of cases. It is important to remember that even in most of these instances, subgroup achievement still improved, although not at the same rate as for white and non-low-income students. Therefore, partly as a result of this pattern of overall improvement, progress on achievement gaps is slow and uneven. Still, it is encouraging to see progress in grade 4 among all subgroups across all three achievement levels—basic, proficient, and advanced—although there is some variation among subgroups.

This report also illustrates how perceived progress on achievement gaps is altered according to the indicator used. While both the percentage proficient and mean score indicators show more instances of gaps narrowing than widening, the mean score indicator gives a less optimistic picture. Which indicator is “better” depends in part on values. Under NCLB, a priority for policymakers is to get lower-performing students to a certain minimal level of proficiency, and for that purpose the percent proficient is a useful indicator. However, if one wants a single indicator that reflects the performance of all students across the spectrum, then the mean test score is more appropriate. There is a danger that the percent proficient indicator may give too rosy a picture of achievement gaps, not only because of the cut score location problem, but also because it only focuses on one point in the student achievement spectrum and does not take into account the performance of students above and below that point. As demonstrated in this study, improvements might be made in student achievement at the advanced or basic levels that do not show up in the percentage proficient measure, thus distorting the picture of progress on achievement gaps.

### Box B. Why gaps change in size depending on where the cut score is set

The relationship between the location of a cut score to the size of a gap was illustrated graphically by Paul Holland (2002). **Figure A** shows the results of the 2000 administration of the NAEP 8<sup>th</sup> grade math test for African American and white students. The test was scored on a scale of 0-500, with the cut score for the basic level of achievement set at 262, proficient at 299, and advanced at 333. The figure shows the percentage of students in each group that scored *at or below* a certain level on NAEP. The x axis is the score, and the y axis is the percentage of students achieving that score or scoring below it. So, about 25% of white students scored at or below 262 (basic)—marked with a dashed vertical line in figure A—while 75% exceeded this score. About 70% of African American students scored at or below 262, while about 30% exceeded this score. Therefore, at the basic level, the achievement gap between African American and white students is about 45 percentage points—quite large.

**Figure A. African American-White achievement gap, NAEP mathematics grade 8, 2000**



Source: Holland, 2002.

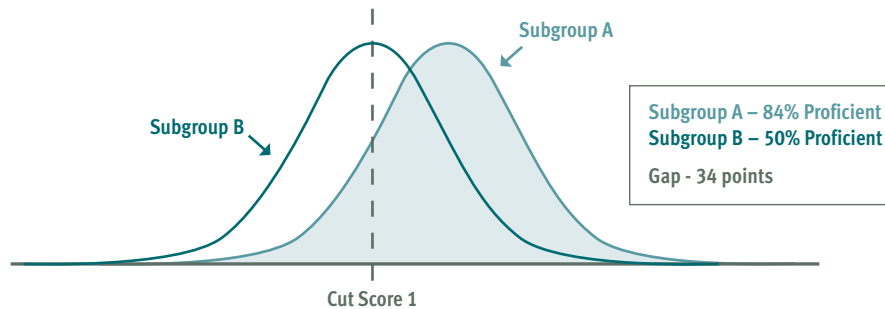
*Continued on next page*

However, the achievement gap picture changes as one moves along the score scale. At the proficient level of 299—marked with a solid vertical line in figure A—the African American-white gap shrinks to about 30 percentage points. As one moves toward the advanced cut score of 333 (shown in the figure as a dotted vertical line), the gap continues to shrink until it reaches about 6 percentage points at the advanced level. The same is true at the low end of the scale, where the gap is also a lot smaller.

As this NAEP example shows, choosing a cut score of 262, 299, or 333 will have a dramatic impact on the apparent size of the achievement gap between African American and white students. The gap is larger at the middle of the NAEP scoring scale than at the extremes.

**Figure B** illustrates this phenomenon in another way. The figure consists of two normal distributions of test scores for two subgroups of students, subgroup A and subgroup B. The figure displays a hypothetical example whereby the initial cut score (cut score 1) is set so that 84% of the students in subgroup A score *at or above* the cut score, compared with 50% of the students in subgroup B. (The areas to the right of the cut score under both curves represent the students who pass.) Therefore, the gap in percentages proficient between the two groups is 34 percentage points.

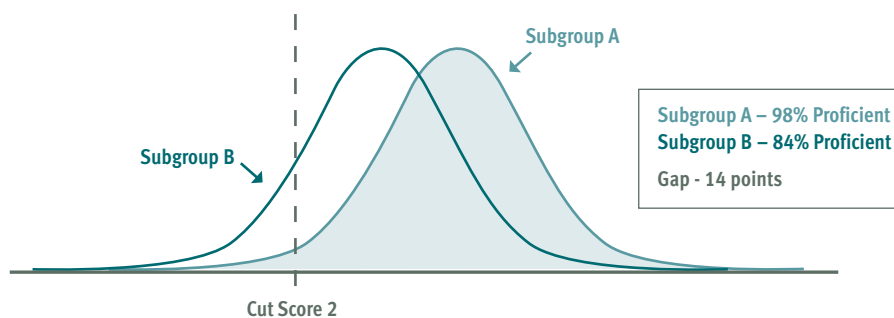
**Figure B. Size of gaps in percentages proficient with a cut score at the mean**



Source: Center on Education Policy.

If a state were to set an easier cut score, represented by cut score 2 in **figure C**, more students would meet or exceed it. At that point, 98% of subgroup A students and 84% of subgroup B students would pass, and the achievement gap would be reduced to 14 percentage points.

**Figure C. Size of gaps in percentages proficient with a lower cut score**



Source: Center on Education Policy.

Therefore, anyone examining trends in achievement gaps must take into account the location of the proficiency cut score.

## References

- Asian American Legal Defense and Education Fund. (2008). *Left in the margins: Asian American students and the No Child Left Behind Act*. Retrieved on September 12, 2009, from [www.aaldef.org/docs/AALDEF\\_LeftintheMargins\\_NCLB.pdf](http://www.aaldef.org/docs/AALDEF_LeftintheMargins_NCLB.pdf)
- Center on Education Policy. (2009a). *State test score trends through 2007-08, part 1: Is the emphasis on proficiency shortchanging higher- and lower-achieving students?* Washington, DC: Author.
- Center on Education Policy. (2009b). *State test score trends through 2007-08, part 2: Is there a plateau effect in test scores?* Washington, DC: Author.
- Center on Education Policy. (2008). *Has student achievement increased since 2002? State test score trends through 2006-07*. Washington, DC: Author.
- Center on Education Policy. (2007). *Answering the question that matters most: Has student achievement increased since No Child Left Behind?* Washington, DC: Author.
- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). *The academic achievement gap in grades three to eight*. Cambridge, MA: National Bureau of Economic Research.
- College Board. (2009a). *2009 college bound seniors: Total group profile report*. Retrieved on August 28, 2009, from <http://professionals.collegeboard.com/profdownload/cbs-2009-national-TOTAL-GROUP.pdf>
- College Board. (2009b). 2009 college-bound seniors are most diverse group ever to take SAT as more minority students prepare for higher education. Retrieved on August 28, 2009, from [www.collegeboard.com/press/releases/206201.html](http://www.collegeboard.com/press/releases/206201.html)
- Dickens, W. T., & Flynn, J. (2006). *Black Americans reduce the racial IQ gap: Evidence from standardization samples*. Retrieved on October 21, 2008, from [www.brookings.edu/articles/2006/10affirmativeaction\\_dickens.aspx](http://www.brookings.edu/articles/2006/10affirmativeaction_dickens.aspx)
- Education Commission of the States. (2008). State subsidizes testing fees for the ACT, SAT, PLAN, PSAT, etc. Retrieved on August 28, 2009, from <http://mb2.ecs.org/reports/Report.aspx?id=1547>
- Fryer, R. G., & Levitt, S. D. (2004). Understanding the black-white test score gap in the first two years of school. National Bureau of Economic Research. Retrieved on August 24, 2009, from [www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content\\_storage\\_01/0000019b/80/1a/c4/dc.pdf](http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1a/c4/dc.pdf)
- Grissmer, D., Flanagan, A., & Williamson, S. (1998). Why did the black-white score gap narrow in the 1970s and 1980s? In C. Jencks & M. Phillips (Eds.), *The black-white test score gap*. Washington, DC: Brookings Institution.
- Harris, D. N., & Herrington, C. D. (2006). Accountability, standards and the growing achievement gap: Lessons from the past half-century. *American Journal of Education*, 112, 209-238.
- Hechinger, J. (2009, August 25). SAT scores fall as gap widens; Asians gain. *Wall Street Journal*.
- Hedges, L. V., & Nowell, A. (1998). Black-white test score convergence since 1965. In C. Jencks & M. Phillips (Eds.), *The black-white test score gap*. Washington, DC: Brookings Institution.
- Holland, P. (2002). Two indicators of change in the gaps between the CDFs of test score distributions. *Journal of Educational and Behavioral Statistics*, 27(1), 3-17.
- Huang, M., & Hauser, R. M. (2000). Convergent trends in black-white test score differentials in the U.S: A correction of Richard Lynn. Center for Demography and Ecology working paper. Retrieved on September 16, 2009, from [www.eric.ed.gov/ERICWebPortal/custom/portlets/recordDetails/detailmini.jsp?\\_nfpb=true&\\_ERICExtSearch\\_SearchValue\\_o=ED460187&ERICExtSearch\\_SearchType\\_o=no&accno=ED460187](http://www.eric.ed.gov/ERICWebPortal/custom/portlets/recordDetails/detailmini.jsp?_nfpb=true&_ERICExtSearch_SearchValue_o=ED460187&ERICExtSearch_SearchType_o=no&accno=ED460187)
- Jencks, C., & Phillips, M. (1998). The black-white test score gap: An introduction. In C. Jencks & M. Phillips (Eds.), *The black-white test score gap*. Washington, DC: Brookings Institution.

- Li, G., & Wang, L. (Eds.). (2008). *Model minority myths revisited: An interdisciplinary approach to demystifying Asian American educational experiences*. Greenwich, CT.: Information Age Publishing.
- National Center for Education Statistics. (2009). *Achievement gaps: How black and white students in public schools perform in mathematics and reading on the National Assessment of Educational Progress*. Retrieved on August 1, 2009, from <http://nces.ed.gov/nationsreportcard/pdf/studies/2009455.pdf>
- National Research Council. (2006). *Multiple origins, uncertain destinies: Hispanics and the American future*, M. Tienda and F. Mitchell (Eds.). Panel on Hispanics in the United States, Committee on Population, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press.
- Rearson, S. F., & Galindo, C. (2008). *The Hispanic-white achievement gap in math and reading in the elementary grades*. Institute for Research on Educational Policy and Practice, Stanford University. Retrieved on August 22, 2009, from [http://irepp.stanford.edu/publications/documents/wp\\_reardon\\_galindo.pdf](http://irepp.stanford.edu/publications/documents/wp_reardon_galindo.pdf)
- Thernstrom, A., & Thernstrom, S. (2004). No excuses. *Closing the racial gap in learning*. New York, NY: Simon and Schuster.
- Zhao, Y., & Qiu, W. (2009, January). How good are the Asians? Refuting four myths about Asian-American academic achievement. Phi Delta Kappan. Retrieved September 15, 2009 from [http://findarticles.com/p/articles/mi\\_6952/is\\_5\\_90/ai\\_n31215064/pg\\_2/?tag=content;col1](http://findarticles.com/p/articles/mi_6952/is_5_90/ai_n31215064/pg_2/?tag=content;col1)

## Credits and Acknowledgments

This report was written by Naomi Chudowsky, Victor Chudowsky, and Nancy Kober, CEP consultants. Naomi and Victor Chudowsky and Jennifer McMurrer conducted research for this study. Sunny Becker, Hilary Campbell, Monica Gribben, Wade Buckland, and Rebecca Dvorak from the Human Resources Research Organization (HumRRO) provided support in data collection and analysis. Diane Stark Rentner, CEP's director of national programs, oversaw the study project for CEP, communicated with the expert panel, and provided advice and assistance for all aspects of the study. Jack Jennings, CEP's president and CEO, provided advice and assistance.

We would like to thank our panel of expert advisors—Laura Hamilton, Eric Hanushek, Frederick Hess, Robert Linn, and W. James Popham—for their invaluable advice. Additionally, we are grateful to the chief state school officers and state assessment personnel for their cooperation in providing information on state testing programs and student achievement data.

Based in Washington, D.C., and founded in January 1995 by Jack Jennings, the Center on Education Policy is a national independent advocate for public education and for more effective public schools. The Center works to help Americans better understand the role of public education in a democracy and the need to improve the academic quality of public schools. We do not represent any special interests. Instead, we help citizens make sense of the conflicting opinions and perceptions about public education and create the conditions that will lead to better public schools.

The Center on Education Policy receives nearly all of its funding from charitable foundations. We are grateful to the William and Flora Hewlett Foundation and the Smith Richardson Foundation for their support of the student achievement study. The George Gund Foundation and the Phi Delta Kappa International Foundation also provide the Center with general support funding that assisted us in this endeavor. The statements made and views expressed are solely the responsibility of the Center.

© Center on Education Policy    October 2009



**Center on Education Policy**

1001 Connecticut Avenue, NW, Suite 522

Washington, D.C. 20036

tel: 202.822.8065

fax: 202.822.6008

e: [cep-dc@cep-dc.org](mailto:cep-dc@cep-dc.org)

w: [www.cep-dc.org](http://www.cep-dc.org)

[www.cep-dc.org](http://www.cep-dc.org)