CENTER ON EDUCATION POLICY CEP

State Test Score Trends Through 2007-08, Part 1

Is the Emphasis on "Proficiency" Shortchanging Higher- and Lower-Achieving Students?

Center on Education Policy

June 2009

CENTER ON EDUCATION POLICY CEP

State Test Score Trends Through 2007-08, Part 1

Is the Emphasis on "Proficiency" Shortchanging Higher- and Lower-Achieving Students?

Center on Education Policy

June 2009

Table of Contents

Main Findings 1

Purpose of the Study, Explanation of Achievement Levels, and Limitations 2

Approach for Analyzing Achievement Trends 4

Trend 1: Student achievement has improved at all achievement levels. 6

Trend 2: Gains tend to be greater at the proficient-and-above level. 8

Trend 3: The percentages of students scoring at the basic level and above have increased in most states. 13

Trend 4: Progress has also been made at the advanced level. 14

Trend 5: More progress has been made in math than in reading. 14

Trend 6: Although achievement has improved at high school, gains have not been as prevalent or large as at other grade levels. 15

Trend 7: An alternative measure of achievement based on average test scores showed similar trends of improved performance. 16

Trend 8: Heavily populated states reflect the national picture. 18

Conclusions about Achievement Trends 19

References 21

Appendix 1. Study Methods and Noteworthy Features 22

Appendix 2. State-by-State Trends at Three Achievement Levels 24

List of Tables, Figures, and Boxes

Box A. A Note on Achievement Levels 3

Table 1. PROFICIENT-AND-ABOVE: Numbers of states showing various trends at this level on state reading and math tests 7

Table 2. BASIC-AND-ABOVE: Numbers of states showing various trends at this level on state reading and math tests 8

Table 3. ADVANCED: Numbers of states showing various trends at this level on state reading and math tests 9

Table 4. Summary of gains and declines as a percentage of all trend lines 10

Figure 1. Average annual percentage point gains 11

Box B. The Difficulty in Comparing Gains at Three Achievement Levels 13

Table 5. Numbers of states showing various trends in effect sizes on state reading and math tests 17

Table 6. Summary of gains and declines in effect size as a percentage of all effect size trend lines 17

Table 7. Trend lines for three achievement levels in the nine most populous states 18

Table A. ELEMENTARY SCHOOL: State–by–state trends on state tests 24

Table B. HIGH SCHOOL: State–by–state trends on state tests 26

Table C. MIDDLE SCHOOL: State–by–state trends on state tests 28

Is the Emphasis on "Proficiency" Shortchanging Higher- and Lower-Achieving Students?

Main Findings

Since the No Child Left Behind Act (NCLB) took effect in 2002, educators have devoted enormous energy to helping students reach the "proficient" level of performance on state tests. The "percentage proficient" has become the vital statistic on the education scorecard because it is used to determine whether districts and schools have made adequate yearly progress (AYP) under NCLB and, if they fall short, whether they must undergo interventions that become increasingly serious over time. But has this intense focus on bringing up students to the proficient level hurt achievement for students at the higher and lower ends of the scale?

For three years, the Center on Education Policy (CEP), an independent nonprofit organization, has studied student achievement trends on the state reading and mathematics tests used for NCLB accountability. This report is the first in a series describing findings from year three of our study. It updates our previous research on overall achievement by adding test results from school year 2007-08, which enables us to discuss trends from 2002 (or the first year after that with comparable data) through 2008. This report also expands on our previous analyses of achievement for students reaching the proficient level (proficient-and-above) by examining trends at two additional achievement levels—students reaching the basic level (basic-and-above) and those attaining the advanced level.

Other reports in this series will consider whether state test scores show a "plateau effect," or leveling off of gains in recent years; will examine trends through 2007-08 in overall performance and achievement gaps for racial-ethnic subgroups and low-income students; will examine recent achievement trends for students with disabilities, English language learners, and male and female students; and will explore the policy implications of our findings about achievement from the other parts of the study.

Here are our main national findings about trends at three achievement levels from the first part of the study:

- **Student achievement in reading and math, as measured by the percentages of students reaching various achievement levels, has generally increased across the board since 2002.** During the years with comparable test data between 2002 and 2008, many more states showed gains than declines at all three achievement levels analyzed.

- **The proficient-and-above level—the target for NCLB purposes—showed the greatest gains.** Out of 300 possible trend lines at the proficient level (two subjects times three grades times 50 states), we had sufficient data to determine 243 trend lines. Of these, 83% showed gains, while 15% showed declines, a better result than at the basic and advanced levels. In addition, the size of the gains was larger, on average, at the proficient level than at the other two achievement levels. This may be partly a statistical phenomenon, as explained in box B. When average test scores go up, the percentage of students at the proficient level tends to grow faster than at the basic and advanced levels because more students' scores tend to be clustered near the proficient level.

- **At the basic-and-above level, most states also made gains.** We found gains in 71% of the trend lines at the basic level and declines in 19%. (The remainder showed no net change.) This means that in most states, the very lowest-performing group—students scoring below the basic level—has shrunk.

- **The advanced level showed more upward than downward trends, as well.** We found gains in 71% of the trend lines at the advanced level and declines in 23%. Moreover, the average gains in percentages of students scoring at the advanced level were often close in size to the gains in percentages proficient. And in elementary math, the average gain at the advanced level was larger. At the advanced level, there were more instances of declines than at the basic and proficient levels, but most of these declines were slight.

- **Although achievement improved at all three grade levels analyzed, there were fewer gains at the high school level than at the elementary or middle school level.** Many more states showed gains than declines among high school students at all achievement levels. But more instances of declining trends emerged at the high school level than at the elementary or middle school level. In high school reading, 15 states had declines in the percentage advanced, nearly equal to the 17 states with gains; however the majority of these declines were slight.

- **We found no strong evidence that NCLB's focus on proficiency is shortchanging students at the advanced or basic levels.** Because the percentage proficient is so crucial to a school's AYP status, there are incentives for educators to make sure they teach what students should know to meet state standards for proficiency and to focus on raising test scores for students who perform slightly or somewhat below proficient—actions that could end up shortchanging higher- or lower-performing students. But state test scores provide little evidence that NCLB is having such an effect. Gains far outnumbered declines at the basic, proficient, and advanced levels of achievement—which suggests that the achievement of higher- and lower-performing students has not been harmed to an obvious extent. Still, gains were more numerous and larger at the proficient than at the basic or advanced levels, which might be interpreted as an effect of NCLB's emphasis on proficiency. However, this is partly due to the statistical phenomenon mentioned above.

Tables summarizing test score trends for each of the 50 states can be found in **appendix 2** at the end of this report.

Purpose of the Study, Explanation of Achievement Levels, and Limitations

This is part 1 of the third annual report by the Center on Education Policy describing trends in state test scores since the No Child Left Behind Act took effect. For three years, CEP has analyzed trends in scores on the state tests in reading (English language arts at the upper grades) and mathematics used for accountability under NCLB. We have also offered possible explanations for these trends. Our objective is to provide complete and accurate data that could inform policy discussions of NCLB and other education issues.

Our previous two studies found that achievement on state tests had risen in most states since 2002 (CEP, 2007; 2008). Last year's study also concluded that trends on the National Assessment of Educational Progress (NAEP) had generally moved in the same direction as

trends on state tests, although gains on NAEP tended to be smaller. Both of those previous studies focused on percentages of students scoring at or above the "proficient" level on state tests because that is the indicator used by most states to determine whether districts and schools have made AYP under NCLB. (A limited number of states use an index for determining AYP that includes the percentage proficient as part of the calculation.) The law requires states to lay out a timeline that will lead to 100% of students scoring proficient by 2014.

This year's study, for the first time, examines trends at the "basic" and "advanced" levels of performance on state tests, as well the proficient level, to obtain a more complete picture of movement across the entire range of student achievement. NCLB requires states to report student test performance in terms of at least three achievement 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.) **Box A** explains some of the challenges involved in analyzing performance at these three levels when there is no common national standard defining the achievement levels and when so much variation exists among state testing systems.

Box A. A Note on Achievement Levels

In this report, we continually refer to the gains made in the percentages of students at three achievement levels: basic and above, proficient and above, and advanced. NCLB requires states to report the percentage of students at each of these levels, but these labels mean different things in different states. Under NCLB, each state has its own tests, which vary in their content, difficulty, question formats, and other aspects. The cut scores that define the various achievement levels on a test, as well as the tests themselves, also vary greatly from state to state, so what a student must know to reach the basic, proficient, or advanced level may be quite different in one state than in another. As a result, one state may have higher percentages of students reaching proficiency than another state, not because its students are learning more but because its tests are easier or its cut scores are lower. In addition, some states have set additional achievement levels beyond the three required by law, and several states use different names for the three main achievement levels (such as "approaches the standard" for basic and "meets the standard" for proficient).

To understand better the range of proportions of students reaching the proficent level across states, we rank-ordered the states by the percentage proficient in elementary math in 2008. The range was huge. In one state, 93.7% of elementary students were proficient in math, while in another state, just 39% were proficient. Similarly, a National Center for Education Statistics study compared state proficiency cut scores to NAEP and "mapped" them onto the NAEP score scale to determine an equivalent NAEP score. The study found that NAEP equivalents of state proficiency cut scores for grade 8 reading ranged from 217 to 278, on a scale of 0-500. All states fell below NAEP's proficiency cut score of 281 (NCES, 2007).

In this report, we discuss the gains and declines in the percentage of students reaching or exceeding the basic, proficient, and advanced levels as defined by each of the 50 states. Readers should keep in mind that there is no common standard for what constitutes performance at these achievement levels. In states with more than three achievement levels or different labels for their levels, we translated the state-defined achievement levels into the three main levels required by NCLB, as explained in the next section.

Because of NCLB's focus on the percentage proficient, concerns have arisen that educators may be responding to federal and state accountability demands by disproportionately targeting their instruction on helping students reach the proficient level. As a result, low-performing students may be receiving instruction that is too difficult for them, and high-performing students may be receiving instruction on content they have already mastered. Some analysts have also expressed concern that teachers may be concentrating their efforts on raising test scores for the so-called "bubble kids" at the expense of higher- and lower-achieving students because doing so will yield a greater payoff in a school's AYP status. (Bubble kids are students who perform slightly or somewhat below the passing score for proficiency.)

If current test-based accountability does cause an inordinate amount of attention to the proficient level in one way or another, one might see rising percentages of students at the proficient-and-above level but stagnant or declining percentages at the advanced or basic-and-above levels (keeping in mind that a decline in the percentage basic-and-above means an increase in the percentage performing below basic). This year's study sought to investigate whether these kinds of undesirable trends had occurred. It builds on work conducted last year by researchers from the Human Resources Research Organization (HumRRO), our contractor for our annual achievement studies (Becker & Campbell, 2009; Gribben, Campbell, & Mathew, 2008).

We recognize that tests are not perfect measures of student achievement. Many educators and researchers have debated the value of state tests as measures of student achievement, given their imprecision, the lack of a common nationwide definition of proficiency, claims that state tests are not sensitive enough to distinguish between good and bad instruction, the possibility that scores on high-stakes tests can be inflated by adjusting instruction to the content of tests, and other dynamics. At the same time, state tests, along with NAEP, are an important source of information on student achievement; indeed, state tests may reflect more closely what is taught in classrooms than NAEP does. CEP and others have addressed these benefits and limitations of tests elsewhere (CEP, 2008; 2007; 2002; Popham, 2007).

Finally, in this or any other study of student achievement it is difficult to sort out the extent to which recent trends have occurred as a result of No Child Left Behind because states, school districts, and schools have simultaneously implemented a variety of different but interconnected policies to raise student achievement since 2002.

Approach for Analyzing Achievement Trends

The main measure of student achievement for this study consists of data from the state tests in reading (or English language arts) and mathematics used for NCLB accountability. Although no large-scale test provides a complete picture of student achievement, we have analyzed state test results because these tests are given to nearly all students in a state, are intended to reflect each state's academic content standards, and are designed to assess whether students have met their states' expectations for performance at a particular grade level.

We carried out this study with advice from a panel of five nationally known experts in educational testing or education policy. We also incorporated several unique features into the design of the study, explained below. More detailed information about study methods and noteworthy features, including the list of expert panel members, appears in **appendix 1** at the end of this report.

- **Data used in the study.** To carry out our achievement studies over the past three years, CEP has compiled a database of state student achievement data from all 50 states, with considerable technical support from HumRRO, our contractor. Data were collected with the cooperation of state education officials, who verified the data's accuracy. Using this database, CEP and HumRRO created a comprehensive data profile for every state, available at www.cep-dc.org. From the 50 states, there were 900 possible achievement trend lines—two subjects, three grade levels, and three achievement levels for each state. For reasons discussed in appendix 1, not all states had sufficient comparable data for all subjects, grades, or achievement levels, so the actual number of trend lines we could analyze was 669—243 at the proficient-and-above level, 236 at advanced, and 190 at basic-and-above. Results were determined by counting the numbers of trend lines showing gains, declines, or no change.

- **Years included in trends.** The achievement trends analyzed for this report cover the years with comparable data between 2002 and 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 earliest starting point because many states did not break out their test results by three or more achievement levels 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). Only those states with 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. Many states have changed their tests in recent years, so the trend lines analyzed in this study often begin after 2002 but no later than 2006.

- **Explanation of achievement categories.** In general, we looked at the percentages of students performing 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 the percentage of students reaching the advanced level, and the percentage basic and above also includes the percentages 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 AYP is determined and is a simpler way to interpret patterns of trends that can become quite complex. If achievement is improving across the board as hoped for, gains should be evident at all three achievement levels—only the percentage of students below the basic level should decrease. In states with more than three achievement levels or different labels for their levels, we translated the state-defined achievement levels into the three main levels required by NCLB, based on state verification of which level represented the proficient level for NCLB accountability. Ohio, for example, has five levels—limited/below basic, basic, proficient, accelerated, and advanced. For our study, we combined data from Ohio's accelerated and advanced categories to form our advanced level, treated basic as our basic level and proficient as proficient, and treated the limited/below basic category as below basic.

- **Averages over three or more years.** To address some of the natural annual fluctuations in test scores that occur for reasons unrelated to students' learning and may yield misleading conclusions, we considered a movement in test results to be a "trend" only if it was based on at least three consecutive years of data. We also determined trends by looking at averages of yearly changes in achievement across three or more years.

- **Size of gains and declines.** Using criteria developed for our previous achievement studies, we defined a moderate-to-large gain or decline as a change of 1.0 percentage point or more, and a slight gain or decline as a change of less than 1.0 percentage point.

The sections that follow explain the achievement trends emerging from this study.

Trend 1: Student achievement has improved at all achievement levels.

Between 2002 (or later in many states) and 2008, many more states showed gains than declines at all three achievement levels analyzed.

PROFICIENT-AND-ABOVE

Table 1 presents the numbers of states that made gains and declines in the percentage of students scoring at the proficient level and above. The number of states with gains in the percentage proficient and above far exceeded the number with declines. For example, in elementary reading, 36 states made gains at the proficient level, while just 7 showed declines. Most of the gains fell into the moderate-to-large category (an average annual change of 1.0 percentage point or greater), while most of the declines were slight (less than 1.0 percentage point per year).

Out of 300 possible trend lines at the proficient level (two subjects times three grades times 50 states), we had sufficient data to determine 243 trend lines. As displayed in the last column of table 1, there were 202 instances of gains of any size, compared with 36 instances of declines. Five trend lines showed no net change (meaning that although the percentage proficient may have gone up or down slightly from one year to the next, the average annual growth over all the years analyzed amounted to 0.0). In 57 instances, not enough years of comparable data through 2008 were available. In percentage terms, of the 243 total trend lines with sufficient data at the proficient-and-above level, 83% represented increases, while 15% showed decreases.

BASIC-AND-ABOVE

Table 2 summarizes state trends at the basic level and above. Again, in both subjects and at all three grades spans analyzed, the number of states with gains far exceeded the number with declines. Of the 190 trend lines we examined, 135 exhibited gains, most of them moderate-to-large. We found just 37 instances of declines, most of them slight. In percentage terms, 71% of the trend lines showed increases at the basic-and-above level, while 19% showed decreases. In other words, in most states the very lowest-performing group—students scoring below the basic level—has shrunk.

One complication is the number of states with insufficient data at the basic level. For example, data were not available for fully half the states at the basic-and-above level in high school math. Some states were unable to provide data at the basic level at all, and for others, there were too few years of data to discern a trend. Furthermore, a number of states did not have a cut score delineating basic performance from below basic, and simply grouped together all students below proficient.

Table 1. PROFICIENT-AND-ABOVE: Numbers of states showing various trends at this level on state reading and math tests

For years between 2002 and 2008 with comparable data

Trend	Reading			Math			Total Trend Lines
	Elementary	Middle	High	Elementary	Middle	High	
States with gains	36	40	25	37	38	26	202
Moderate-to-large gain	*27*	*29*	*17*	*27*	*34*	*20*	*154*
Slight gain	*9*	*11*	*8*	*10*	*4*	*6*	*48*
States with no change	1	0	0	1	2	1	5
States with declines	7	3	10	6	3	7	36
Moderate-to-large decline	*2*	*2*	*5*	*2*	*1*	*3*	*15*
Slight decline	*5*	*1*	*5*	*4*	*2*	*4*	*21*
Trend data insufficient or unavailable	6	7	15	6	7	16	57

Table reads: Between 2002 and 2008, 36 states showed average annual gains in the percentage of elementary students scoring at the proficient level and above on state reading tests. Of these 36 states, 27 had moderate-to-large average annual gains, and 9 had slight gains. Seven states experienced average annual declines in the percentage of elementary students reaching the proficient level in reading. Six states did not have three or more years of comparable data through 2007-08, so a trend could not be determined.

Note: A "moderate-to-large" gain or decline refers to an average annual change of 1.0 or more percentage points. A "slight" gain or decline refers to an average annual change of less than 1.0 percentage point. "No change" means an average annual change of 0.0.

Note: Trends shown are for students in grade 4 (elementary), grade 8 (middle), and grade 10 (high school) in states with sufficient data at these grades. In other states, an adjacent grade was selected according to a consistent set of criteria.

ADVANCED

Table 3 displays trends at the advanced level. The number of states with gains far outstripped the number with declines in all subjects and grades, except in high school reading, where 17 states had gains while 15 had declines. At the advanced level, we found 167 instances of gains of any size and 54 instances of declines. Overall, 71% of the trend lines were increases, while 23% were decreases. Most of the increases were moderate-to-large, but most of the decreases were slight.

POSSIBLE EXPLANATIONS

There are several possible explanations for the upward trends in student achievement. The most hopeful explanation is that students are learning more and consequently are performing better on state tests. Other possible explanations, though not the only ones, are that districts and schools are reallocating instructional time and effort toward reading and math, the subjects

Table 2. BASIC-AND-ABOVE: Numbers of states showing various trends at this level on state reading and math tests

For years between 2002 and 2008 with comparable data

Trend	Reading			Math			Total Trend Lines
	Elementary	Middle	High	Elementary	Middle	High	
States with gains	21	24	17	28	29	16	135
Moderate-to-large gain	*11*	*14*	*11*	*14*	*22*	*10*	*82*
Slight gain	*10*	*10*	*6*	*14*	*7*	*6*	*53*
States with no change	3	6	3	1	2	3	18
States with declines	10	4	7	6	4	6	37
Moderate-to-large decline	*3*	*0*	*3*	*2*	*1*	*1*	*10*
Slight decline	*7*	*4*	*4*	*4*	*3*	*5*	*27*
Trend data insufficient or unavailable	16	16	23	15	15	25	110

Table reads: Between 2002 and 2008, 21 states showed average annual gains in the percentage of elementary students scoring at the basic level and above on state reading tests. Of these 21 states, 11 had moderate-to-large average annual gains, and 10 had slight gains. Ten states experienced average annual declines in the percentage of elementary students reaching the basic level in reading. Sixteen states did not have three or more years of comparable data through 2007-08, so a trend could not be determined.

Note: A "moderate-to-large" gain or decline refers to an average annual change of 1.0 or more percentage points. A "slight" gain or decline refers to an average annual change of less than 1.0 percentage point. "No change" means an average annual change of 0.0.

Note: Trends shown are for students in grade 4 (elementary), grade 8 (middle), and grade 10 (high school) in states with sufficient data at these grades. In other states, an adjacent grade was selected according to a consistent set of criteria.

tested for NCLB accountability, or that states have made subtle changes in the difficulty or scoring of their tests. A cumulative effect of test-focused instruction may also be a factor. In the first year a test is administered, educators may be surprised initially by what is included on the tests, but they gradually realize the material likely to be covered on subsequent tests and align instruction more accurately with what is assessed. Of course, a downside of this is that some important content that is not tested may be ignored, leading to a narrowing of curricula. These explanations are discussed more fully in our previous reports and elsewhere (see CEP, 2007, pp. 41-42; CEP, 2008, pp. 41-42; and Hamilton et al., 2007)

Table 3. ADVANCED: Numbers of states showing various trends at this level on state reading and math tests

For years between 2002 and 2008 with comparable data

Trend	Reading			Math			Total Trend Lines
	Elementary	Middle	High	Elementary	Middle	High	
States with gains	28	30	17	35	37	20	167
Moderate-to-large gain	*19*	*20*	*9*	*29*	*32*	*12*	*121*
Slight gain	*9*	*10*	*8*	*6*	*5*	*8*	*46*
States with no change	1	3	2	3	4	2	15
States with declines	14	9	15	5	1	10	54
Moderate-to-large decline	*4*	*4*	*5*	*1*	*0*	*4*	*18*
Slight decline	*10*	*5*	*10*	*4*	*1*	*6*	*36*
Trend data insufficient or unavailable	7	8	16	7	8	18	64

Table reads: Between 2002 and 2008, 28 states showed average annual gains in the percentage of elementary students scoring at the advanced level on state reading tests. Of these 28 states, 19 had moderate-to-large average annual gains, and 9 had slight gains. Fourteen states experienced average annual declines in the percentage of elementary students reaching the advanced level in reading. Seven states did not have three or more years of comparable data through 2007-08, so a trend could not be determined.

Note: A "moderate-to-large" gain or decline refers to an average annual change of 1.0 or more percentage points. A "slight" gain or decline refers to an average annual change of less than 1.0 percentage point. "No change" means an average annual change of 0.0.

Note: Trends shown are for students in grade 4 (elementary), grade 8 (middle), and grade 10 (high school) in states with sufficient data at these grades. In other states, an adjacent grade was selected according to a consistent set of criteria.

Trend 2: Gains tend to be greater at the proficient-and-above level.

In terms of both the percentage of gains and declines and the size of the gains, the greatest increases occurred in the percentages of students reaching proficiency.

PERCENTAGES OF GAINS AND DECLINES

Table 4 summarizes gains and declines in trend lines as a percentage of all trend lines for which data were available; it is based on tables 1, 2 and 3. In elementary reading, for example, 62% of states with sufficient data had gains at the basic-and-above level, while 29% had declines. The reason the percentages of gains and declines do not total 100% is that some states with sufficient data exhibited no net change. The "percentage across grades and subjects" column in

Table 4. Summary of gains and declines as a percentage of all trend lines

For years between 2002 and 2008 with comparable data

Achievement Level & Trend	Reading			Math			Percentage Across Grades & Subjects
	Elementary	Middle	High	Elementary	Middle	High	
Basic & above gains	62%	71%	63%	80%	83%	64%	71%
Basic & above declines	29%	12%	26%	17%	11%	24%	19%
Proficient & above gains	82%	93%	71%	84%	88%	76%	83%
Proficient & above declines	16%	7%	29%	14%	7%	21%	15%
Advanced gains	65%	71%	50%	81%	88%	63%	71%
Advanced declines	33%	21%	44%	12%	2%	31%	23%

Table reads: Between 2002 and 2008, the percentage of students performing at the basic level and above on state elementary reading tests increased in 62% of the states with sufficient years of comparable data and declined in 29%. During this same period, the percentage of students performing at the proficient level and above on state elementary reading tests increased in 82% of the states with sufficient data and declined in 16%.

Note: Trends shown are for students in grade 4 (elementary), grade 8 (middle), and grade 10 (high school) in states with sufficient data at these grades. In other states, an adjacent grade was selected according to a consistent set of criteria.

table 4 presents the percentage of gains or declines across all states with data, in both subjects and at three grade levels. For example, of all the trend lines at the basic-and-above level nationwide, 71% showed gains, while 19% showed declines.

The highest percentage of gains occurred at the proficient-and-above level. Of all trend lines at this level, 83% showed increases, higher than the percentage of gains at the basic (71%) or advanced (71%) levels.

SIZE OF PERCENTAGE POINT GAINS

We also analyzed the size of the gains and declines in the percentages of students performing at each achievement level. We determined the size of a change by calculating the average annual percentage point gain or decline in a particular state for each subject and grade level. Then we arrived at a national figure by averaging these averages across all 50 states. The average annual percentage point gain or decline 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, 51% of high school students scored at or above proficient in reading in 2008, compared with 37% in 2002—a difference of 14 percentage points. The average annual gain is therefore 14 percentage points divided by six years, or 2.3 percentage points per year.

Average percentage point gains were usually largest at the proficient level, as shown in **figure 1.** In elementary reading, the average annual gain across all states was 1.3 percentage points at the proficient-and-above level, higher than the average gain of 0.5 percentage point at basic-and-above or 1.0 percentage point at advanced. Similarly, in middle school math, the average annual gain was 2.3 percentage points at proficient-and-above, compared with 1.4 at basic-and-above and 1.9 at advanced. This pattern emerged in all subjects and grades except elementary math, where the largest average percentage point gain (1.8 percentage points) occurred at the advanced level—greater than the gains of 1.7 at the proficient level and 0.8 at the basic level.

POSSIBLE EXPLANATIONS

Several factors may explain why more states had gains at the proficient level than at the basic or advanced level and why there were greater percentage point gains at the proficient level.

- **When average test scores increase, a greater percentage of students tend to move into the proficient level than the other two achievement levels.** This happens because in a typical population of students, more students score at or around the proficient level than score at or around the basic or advanced levels. Therefore, when average test scores improve over time, a larger number, and a higher percentage, of students cross the threshold from basic to proficient. The number and percentage of students crossing the threshold from below basic to basic, or from proficient to advanced, tends to be smaller. A more detailed explanation is provided in **box B**.

- **Teachers have incentives to focus on bringing up students to proficiency.** NCLB could be encouraging teachers and administrators to concentrate their efforts on reaching the proficient level in two ways: by focusing on academic content that is most closely aligned to the proficient level on tests, or by devoting more attention to "bubble kids,"

Figure 1. Average annual percentage point gains

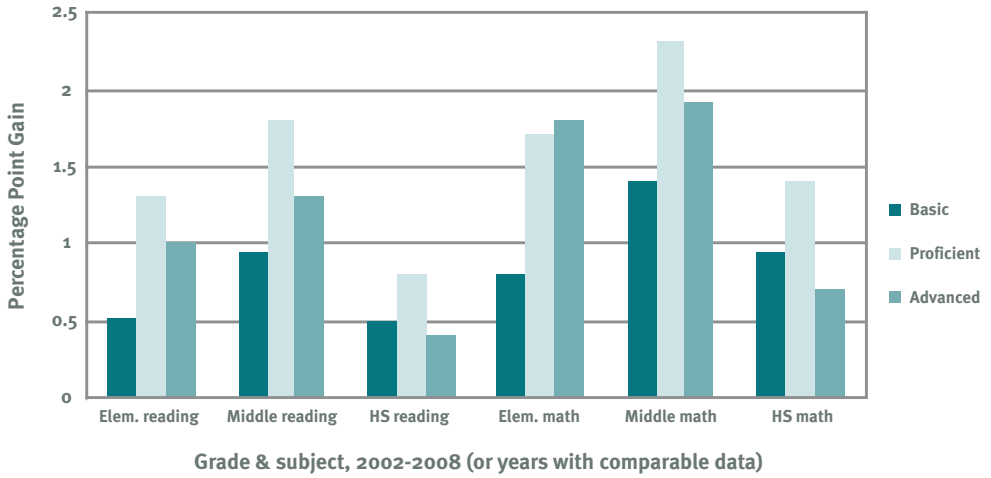

Figure reads: In elementary reading, the percentage of students across all 50 states scoring at the basic level and above increased by an average of 0.5 percentage point per year between 2002 and 2008. This was lower than the average annual gain in elementary reading of 1.3 percentage points at the proficient level and above and 1.0 points at the advanced level.

those students who score slightly or somewhat below the proficiency threshold (Booher-Jennings, 2005; Reback, 2006; Hamilton et al., 2007; Loveless et al., 2008). In a 2007 study, Neal and Schanzenbach examined test score distributions for a group of students before and after implementation of test-based accountability in Chicago in 1996, and then again in 2002. They found that scores improved at the middle of the score distribution but were unchanged for students at the bottom and increased only slightly for students at the top. However, there is not consensus on this issue among researchers. Springer (2008) used a database of test scores of 300,000 students in one western state, focusing on schools that had not made adequate yearly progress under NCLB. In these schools, the greatest test score gains were made by the lowest-performing students rather than those near the proficiency threshold.

- **High percentages of students are already scoring at the basic level.** Gains may be more numerous and larger at the proficient level than at the basic level because in most states, the benchmark for the basic level of achievement is low enough that the vast majority of students have already met it in most subjects and grades. A "ceiling effect" occurs, whereby it is difficult to increase this percentage much further because any gains must be achieved by the small percentage of students below basic, who are likely to have the most serious learning challenges. For example, in middle school math in 2005, the average percentage of students scoring basic and above across all states was 87.4%. Some states had upwards of 95% of students at basic and above.

- **Tests may be less sensitive to improvements in advanced skills.** Most standardized tests, including some state tests, do a better job of differentiating student performance around the middle of the achievement distribution than at either the high or low end. For example, a test may have fewer difficult questions that discriminate between students who tend to score high versus low on the test overall. Tests have larger measurement error at both extremes of the test score distribution. The tests tend to be targeted to provide the most information, which can be thought of as the most precise measure, in the middle of the distribution, which is usually close to the proficient cut score. Also, state tests may be most closely aligned to the content of middle-level classes. For example, if many of the advanced students in a school are enrolled in 8th grade algebra, a state 8th grade math test that covers relatively little algebra might not be very sensitive to learning gains among these students because the content of their course emphasizes material that is not well represented on the test. These students may show few gains on the state test despite knowing a great deal of algebra.

- **Schools and teachers may not be well-equipped to effectively reach the lowest- or highest-performing students.** Instructional methods that work well with students in the broad middle range of performance may not be as effective with lower- or higher-achieving students. Special knowledge and experience may be required to make significant progress with students at either end of the educational spectrum.

Box B. The Difficulty in Comparing Gains at Three Achievement Levels

A key indicator of student achievement under current accountability systems is the percentage of students reaching certain state-defined achievement levels, such as basic, proficient, and advanced. It is difficult to compare gains in the percentages of students reaching these achievement levels because of a statistical issue: the cut scores for each achievement level are located where the density of the distribution is substantially different. Usually, the distribution of test scores approximates the shape of a normal curve. More students are clustered around the proficient level than the advanced or basic levels, so if the average test score increases, a greater number and higher percentage of students will move into the proficient level than the other two achievement levels.

To see if this issue explained some of our results, we calculated the average percentage of students across all 50 states at each achievement level in the first year of the trend line, by grade and subject. For example, for high school math, these averages came to 76.7% at basic and above, 56.4% at proficient and above, and 18.5% advanced. Then we looked at how large of an increase in average test scores would be necessary to increase the percentage of students at each achievement level by five percentage points. To accomplish this gain, the average test score would have to increase by 0.28 standard deviations (SD) at the basic level, by 0.13 SD at the proficient level, and by 0.18 SD at the advanced level. In other words, a far greater increase in average test scores is necessary to increase the percentage of students at the basic level by five points than at the other two achievement levels. Improving the percentage at the proficient level would take the smallest increase in average test scores.

The same pattern exists with the other grade/subject combinations, with the exception of elementary math, where the increase in average test scores necessary to increase both the percentage proficient and percentage advanced was identical (0.15 SD) but still smaller than the increase needed to raise the percentage basic by five points (0.29 SD).

Therefore, it is more difficult to show improvements at the basic achievement level than at the proficient level, simply because of this statistical issue. (This leaves aside other difficulties in moving more students into the basic-and-above achievement level, such as the fact the students who are below the basic level may have the most serious learning challenges). It is also more difficult to produce gains in percentages of students at the advanced level than at the proficient level. Our finding of larger percentage point gains at the proficient level is partly attributable to this statistical phenomenon.

Trend 3: The percentages of students scoring at the basic level and above have increased in most states.

As noted above and in table 2, we found more gains than declines at the basic-and-above level. Even so, several states did show downward trends. We delved into the data from the 10 states with declines in elementary reading and found no obvious explanation for this trend. Most of the declines at the basic-and-above level were slight. None of the states had a moderate-to-large decline in the percentage reaching basic coupled with a moderate-to-large gain in the percentage reaching proficient. In other words, none had a strong pattern that would suggest low-achieving students were being shortchanged while students closer to proficiency were receiving undue attention. However, four states did show a slight decrease in the percentage basic coupled with an increase in the percentage proficient.

Trend 4: Progress has also been made at the advanced level.

Our analysis revealed many more upward than downward trends in the percentage of students at the advanced level (table 3). Moreover, in elementary and middle school math, the percentages of gains at the advanced level were close or equal to those at the proficient level (table 4). We also found average annual percentage point gains at the advanced level that were close in size to those at the proficient level (figure 1).

Based on NAEP data, some researchers have expressed concern that the performance of high-achieving students has languished since the inception of NCLB, with its focus on the proficient level (Loveless et al., 2008). We cannot directly dispute these findings, as our methodology is different and we are using state tests as measures instead of NAEP, which has more demanding thresholds for basic, proficient, and advanced performance than state tests do. Our study focused on the increase in the percentage of students moving into the advanced level on state tests. We did not look at changes in achievement among students who were already performing at the advanced level.

Our findings about the preponderance of gains at the advanced level generally confirm the trends uncovered by Gribben, Campbell, and Mathew (2008) and point to a significant movement of students from proficient to advanced. If teachers are indeed targeting their instruction at the proficient level, the effect is not so extreme as to thwart growth at the advanced level. Again, as explained in the section on trends at the proficient level and box B, the use of percentages of students reaching various achievement levels as indicators of progress may exaggerate gains in the percentage proficient relative to the percentage advanced.

Amid these positive signs, we also found some evidence of declines at the advanced level that bear watching in the future. In elementary reading, 14 states experienced declines in the percentage advanced—10 with slight declines, and 4 with moderate-to-large declines. The high school level did not fare as well as the elementary or middle school level in terms of trend lines with gains and declines at the advanced level. In both high school reading and math, roughly one-fifth to one-third of all states showed declines at the advanced level. In high school reading, the number of states with declines in the percentage advanced was close to the number with gains (15 states with declines versus 17 with gains). In high school math, 10 states posted declines at the advanced level (again, most of these were slight). The high school grade analyzed had the smallest average point gain in the percentage of students reaching the advanced level. And unlike the other two grade levels, the high school grade made less progress in moving students into the advanced level than into the basic level (figure 1). These trends can be put into the context of generally less positive results at the high school grades, discussed below.

Trend 5: More progress has been made in math than in reading.

In general, states showed more gains in math than in reading at all three achievement levels—basic-and-above, proficient-and-above, and advanced (table 4). In addition, the average annual percentage point gains across all states were larger in math than in reading (figure 1). NAEP data confirm this general finding. Results from the NAEP long-term trend assessments indicate that score gains since the 1970s have been much greater in math than in reading (NCES, n.d.). CEP's 2008 achievement study, which analyzed state-level results from a different set of NAEP assessments, the state NAEP, also found more states with gains in math than in reading. These disparities in achievement growth between math and reading cannot be attributed simply to differences in the difficulty of state tests in these subjects.

One possible explanation for this trend may relate to significant differences between teaching students to acquire math skills and teaching them to comprehend what they read. Math skills are more discrete and are based on certain rules that can be systematically taught to students, which may make them easier to measure with standardized tests. Learning to read is a cumulative process that sometimes takes place over several years and requires students to apply a broad range of skills and knowledge in an integrated way. Teachers may also know more about effective ways to teach math to students who are having difficulty.

Another explanation for this trend may be statistical. We took a closer look at the average percentages, across all states with comparable data, of students reaching proficiency in math and reading. We found that at the middle and high school levels, percentages proficient were lower in math, but growth was greater. In other words, there is more room for growth in math at these levels.

Trend 6: Although achievement has improved at high school, gains have not been as prevalent or large as at other grade levels.

High school results were somewhat mixed. On the positive side, many more states showed gains than declines among high school students for all achievement level-subject combinations (tables 1, 2, and 3). Overall, the percentage of high school students reaching or exceeding the three achievement levels has increased. However, the downside comes in the comparison between high school and the two lower grade levels. As table 4 illustrates, a smaller percentage of states in almost every category (reading and math at the three achievement levels) made gains at the high school level than at the lower two grade levels, with the exception of reading at the basic level. At the advanced level, only a little more than half of the states with trend data made progress in high school reading. Improvement was slowest as well; average annual percentage point gains tended to be smallest at the high school level (figure 1).

Our finding that the gains were less pronounced at the high school level is not surprising. Evidence from other studies points to some troubling trends at this level. Generally speaking, test scores have gone up nationwide, but in many states high school achievement has run counter to that trend. Other studies have found greater increases in state test scores at the elementary and middle grades than at high school since the inception of NCLB (Education Trust, 2005). National data from the long-term NAEP assessments show gains among 9- and 13-year-old students in both reading and math over the past three decades but stagnant scores for high school students since 1971 in reading and since 1978 in math (National Center on Education Statistics, n.d.).

Several factors might explain the divergence in performance between students at the lower and higher grades:

- **High school students may be less engaged in school and less motivated to do well.** A sizable body of research from psychology indicates that as students progress through the grades and enter high school, their self-motivation to perform well academically tends to decrease, as does their desire to please their teachers (for a review of this literature, see Lepper & Henderlong, 2000). The National Research Council (2003) has also summarized and synthesized research on the factors leading to disengagement and alienation among high school students. A narrower explanation is that high school students are less motivated to perform well on state tests, particularly if the tests do not count

toward their graduation requirements or are not used to determine course grades. Even high-achieving students may choose not to put forth their best effort on a state test that has few or no consequences for them individually.

- **High school students are subject to more outside influences.** Researchers have also examined family and peer influences and the amount of time high school students spend working at part-time jobs (Quirk, Keith, & Quirk, 2001; Steinberg, 1988).

- **High schools receive fewer resources from federal programs.** The majority of federal Title I funds for low-achieving students in low-income areas goes to elementary schools; high schools receive fewer Title I funds.

- **High school course content may be less well-aligned with tested material.** At the high school level, the content of courses students take during the tested year may be less well matched with the content of the actual state test. For example, 11th graders may be taking algebra II, precalculus, trigonometry, Advanced Placement courses, or other courses that do not match the content of the state test very well. However, end-of-course assessments, which some states use as their high school tests for NCLB purposes, should be highly aligned to course content.

Trend 7: An alternative measure of achievement based on average test scores showed similar trends of improved performance.

Effect sizes are an indicator of achievement based on average, or mean, test scores. Because effect sizes do not depend on the cut score a state has set for proficient performance, they are not susceptible to some of the limitations of percentages proficient. As we did in our 2007 and 2008 reports, we looked at effect size as a second indicator of achievement.

Effect size data were unavailable in several states, especially at the high school level. In the states that did provide these data, the effect size trends were generally consistent with the trends in percentage proficient described above. The instances of gains in effect sizes far outnumbered instances of declines (see **table 5**). In middle school math, the subject with the highest ratio of gains to declines, 32 states had gains in effect sizes and 3 had declines. Middle school reading also showed strong improvement: 30 states had gains and just 3 had declines. In elementary reading and math, many more states experienced gains in effect size than declines. As with the percentage proficient analyses, the number of states with gains at high school was lower than for the elementary and middle school grades.

Statistics on the percentage of trend lines showing increases in effect sizes (see **table 6**) lend support to our general finding that achievement on state tests has improved. Out of a total of 197 effect size trend lines with sufficient data, 77% demonstrated gains and just 15% demonstrated declines. The 77% of trend lines with effect size gains is close to the proportion of trend lines with gains at the proficient-and-above level (83%, see table 4) and higher than the proportion with gains at the basic-and-above (73%) and advanced (71%) levels.

Table 5. Numbers of states showing various trends in effect sizes on state reading and math tests

For years between 2002 and 2008 with comparable data

Trend	Reading			Math			Total Trend Lines
	Elementary	Middle	High	Elementary	Middle	High	
States with gains	25	30	17	31	32	17	152
Moderate-to-large gain	*22*	*24*	*12*	*27*	*31*	*12*	*128*
Slight gain	*3*	*6*	*5*	*4*	*1*	*5*	*24*
States with no change	5	3	2	1	1	3	15
States with declines	6	3	8	4	3	6	30
Moderate-to-large decline	*3*	*3*	*6*	*3*	*1*	*4*	*20*
Slight decline	*3*	*0*	*2*	*1*	*2*	*2*	*10*
Trend data insufficient or unavailable	14	14	23	14	14	24	103

Table reads: Between 2002 and 2008, 25 states showed average annual gains in effect size on state reading tests at the elementary level. Of these 25 states, 22 had moderate-to-large average annual gains, and 3 had slight gains. Six states experienced average annual declines in effect size in elementary reading. Fourteen states did not have three or more years of comparable data through 2007-08, so a trend could not be determined.

Note: A "moderate-to-large" gain or decline in effect size refers to an average annual change of 0.02 standard deviation units or more. A "slight" gain or decline refers to an average annual change of less than 0.02. "No change" means an average annual change of 0.00.

Note: Trends shown are for students in grade 4 (elementary), grade 8 (middle), and grade 10 (high school) in states with sufficient data at these grades. In other states, an adjacent grade was selected according to a consistent set of criteria.

Table 6. Summary of gains and declines in effect size as a percentage of all effect size trend lines

For years between 2002 and 2008 with comparable data

Trend	Reading			Math			Percentage Across Grades & Subjects
	Elementary	Middle	High	Elementary	Middle	High	
Gains	69%	83%	63%	86%	89%	65%	77%
Declines	17%	8%	30%	11%	8%	23%	15%

Table reads: Between 2002 and 2008, 69% of the states with sufficient years of comparable data showed gains on elementary reading tests according to effect sizes, while 17% of these states showed declines.

Note: Trends shown are for students in grade 4 (elementary), grade 8 (middle), and grade 10 (high school) in states with sufficient data at these grades. In other states, an adjacent grade was selected according to a consistent set of criteria.

Trend 8: Heavily populated states reflect the national picture.

We did not use any weighting system that would account for the population of students in each state; we treated the population of students taking California's tests the same as the population taking Wyoming's tests. But if students in less populous states are doing better than those in heavily populated states, we may not be getting an accurate picture of what is happening with student achievement nationwide.

To see if this was the case, we compiled trend lines from the nine most populous states, which together account for about half of the student enrollment nationwide (these include California, Texas, New York, Florida, Illinois, Ohio, Pennsylvania, Michigan, and Georgia). As seen in **table 7**, the picture of gains and declines is similar to the overall results presented above. Gains far outnumbered declines at each achievement level, and gains were more numerous at the proficient level than the advanced or basic levels. As in the overall results, there were more gains in math than reading.

Table 7. Trend lines for three achievement levels in the nine most populous states

For years between 2002 and 2008 with comparable data

Trend	Reading			Math		
	Basic & Above	Proficient & Above	Advanced	Basic & Above	Proficient & Above	Advanced
Moderate-to-large gain	8	18	14	10	19	17
Slight gain	6	5	4	7	1	4
Moderate-to-large decline	0	1	1	1	1	1
Slight decline	5	0	5	1	2	1
Trend data insufficient or unavailable	8	3	3	8	4	4

Table reads: Of the trends analyzed at three achievement levels in the nine most populous states, eight trends showed moderate-to-large gains in reading at the basic level and above, and six trends showed slight gains. No trends showed moderate-to-large declines in reading, and five trends showed slight declines. In eight cases, a trend could not be determined because the states did not have three or more years of comparable data through 2007-08.

Note: A "moderate-to-large" gain or decline refers to an average annual change of 1.0 or more percentage points. A "slight" gain or decline refers to an average annual change of less than 1.0 percentage points. "No change" means an average annual change of 0.0.

We also broke down the trend lines for these nine states in a different way, by grade level. As was the case nationally, the results were better at the elementary and middle school levels than at high school. Across the two subjects in the elementary grade, 40 trend lines showed gains and 8 showed declines. For middle school, the comparable figures were 46 trend lines with gains and only 2 with declines. At the high school level, 27 trend lines showed gains and 9 showed declines.

Conclusions about Achievement Trends

In this study, we sought to present a detailed analysis of state test score trends based on accurate data to inform policy discussions of NCLB and other education issues. We found that overall achievement, as measured by state test scores, has continued to rise and that the greatest gains have been made at the proficient level. We did not find strong evidence that NCLB's emphasis on the proficient level was occurring at the expense of high- and low-achieving students. If accountability policies were indeed shortchanging high- and low-achieving students, we would expect to see stagnation or decline at the basic and advanced levels. Instead, the percentages of students scoring at the basic-and-above and advanced levels have increased much more often than they have decreased, especially in the elementary and middle grades.

Gains were larger at the proficient level, but to some extent this was due to a statistical phenomenon—when mean test scores go up, the percentage of students at the proficient level grows faster than at the advanced level because more students were already scoring near the proficient level. After analyzing the percentages of students performing at the three achievement levels during the initial year of the trend lines for state test scores, we determined that a greater increase in average test scores would be necessary to show improvement at the basic or advanced levels than at the proficient level. Other factors may also explain the greater increases at the proficient level, such as the limits on growth in the basic level and the possibility that tests are more sensitive in measuring improvements at the proficient than at the advanced level. Finally, our findings also raise the important question of why the extent and magnitude of gains in high school do not match those in the lower grades.

See Appendix 2 and CEP's Web Site for More Detailed State-by-State Information

This report gives an overview of student achievement trends across the 50 states. The major trends for individual states are displayed in three tables, one for each grade level, in appendix 2 to this report.

Additional detailed state-by-state information is available on CEP's Web site, www.cep-dc.org. There, readers can find a wealth of information for specific states:

- A **state profile** for each of the 50 states that contains the following information—
    - Narrative bullet points summarizing the main achievement trends in that state at the basic, proficient, and advanced levels
    - Background information about characteristics of the state's testing program, changes in its testing program over time, and limitations of the state's test data
    - Figures and tables with detailed data illustrating trends at three grade levels (elementary, middle, and high school), two subjects (reading and mathematics), and three achievement levels, including discrete percentages ("basic" alone instead of "basic and above" and "proficient" alone instead of "proficient and above")
    - Effect size data with mean test scores and standard deviations
    - A glossary of key terms and a box with caveats and explanations

- Tables displaying **the detailed data used to develop tables 1 through 3** in this report, including state abbreviations indicating which states exhibited which trends
- Versions of the tables in appendix 2 showing **average percentage point gains and declines** for each of the 50 states in numeric terms
- Tables with **effect size trends** for each of the 50 states
- Tables showing the **years of comparable data** for each state
- A table showing the **grade levels** analyzed in states where sufficient comparable data were unavailable for the default grades 4, 8, and 10
- **Raw test data** collected from the states in Excel files (available in August 2009)

References

Becker, D. E., & Campbell, H. L. (2009, April). Left behind or rising with the tide? Trends in basic+ achievement on 50 state assessments. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Booher-Jennings, J. (2005). Below the bubble: Educational triage and the Texas accountability system. *American Educational Research Journal (42)*2, 231-268.

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.

Center on Education Policy. (2002). Teaching to the test: The good, the bad, and who's responsible. *TestTalk for Leaders (1)*. Washington, DC: Author.

Editorial Projects in Education. (2004, September 21). High school reform. *Education Week.* Retrieved April 14, 2009, from www.edweek.org/rc/issues/high-school-reform/

Education Trust. (2005). *Stalled in secondary: A look at student achievement since the No Child Left Behind Act.* Washington, DC: Author.

Gribben, M.L., Campbell, H.L., & Mathew, J. (2008, March). *Are advanced students advancing? Examining achievement trends beyond proficiency.* Paper presented at the annual meeting of the American Educational Research Association, New York.

Hamilton, L. S., Stecher, B. M., Marsh, J. A., McCombs, J. S., Robyn, A., Russell, J. L., Naftel, S., & Barney, H. (2007). *Standards-based accountability under No Child Left Behind.* Santa Monica, CA: RAND Corp.

Lepper, M. R., & Henderlong, J. (2000). Turning "play" into "work" and "work" into "play": 25 years of research on intrinsic versus extrinsic motivation. In C. Sansone & J.M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance.* San Diego, CA: Academic Press.

Loveless, T., Farkas, S., & Duffett, A. (2008). *High-achieving students in the era of NCLB.* Washington, DC: Thomas S. Fordham Institute. Retrieved on April 23, 2009, from www.edexcellence.net/doc/20080618_high_achievers.pdf

National Center for Education Statistics. (n.d.). NAEP data explorer. http://nces.ed.gov/nationsreportcard/lttdata/

National Center for Education Statistics (2007). *Mapping 2005 state proficiency standards onto the NAEP scales.* NCES 2007-482. Washington, DC: U.S. Department of Education.

National Research Council. (2003). *Engaging schools: Fostering high school students' motivation to learn.* Washington DC: National Academies Press. Retrieved April 14, 2009, from www.nap.edu/openbook.php?record_id=10421&page=R1

Neal, D. & Schanzenbach, D. W. (2007). Left behind by design: Proficiency counts and test-based accountability. Unpublished manuscript. Retrieved April 14, 2009, from /www.aei.org/doclib/20070716_nealschanzenbachpaper.pdf

Popham, W. J. (2007, October). Instructional insensitivity of tests: Accountability's dire drawback. *Phi Delta Kappan, (89)*2, 146-150.

Quirk, K. J., Keith, T. Z., & Quirk, J. T. (2001). Employment during high school and student achievement: Longitudinal analysis of national data. *Journal of Educational Research (95)*, 4-10.

Reback, R. (2006). *Teaching to the rating: School accountability and the distribution of student achievement.* New York: Barnard College, Columbia University.

Springer, M. G. (2008, Winter). Do schools practice educational triage? *Education Next, (8)*1, 74-79. Retrieved May 12, 2009, from http://media.hoover.org/documents/ednext_20081_74.pdf

Steinberg, L. (1988). *Noninstructional influences on high school student achievement: The contributions of parents, peers, extracurricular activities, and part-time work.* Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement. Retrieved April 14, 2009, from www.eric.ed.gov/ERICWebPortal/custom/portlets/recordDetails/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED307509&ERICExtSearch_SearchType_0=no&accno=ED307509

Appendix 1. Study Methods and Noteworthy Features

We took several steps to overcome some of the common problems with state test data, improve the accuracy of the data analyzed, and address factors that could lead to misleading conclusions about achievement trends. Several study features and methods are particularly noteworthy:

- **All 50 states participated.** As they did the past two years, all states participated, making this the most comprehensive study of achievement trends since NCLB was enacted. (The District of Columbia chose not to participate in any year of the study.) We greatly appreciate the cooperation we received from state education personnel in all states.

- **States verified the accuracy of their data.** CEP and HumRRO, our contractor, gathered the relevant test data from state Web sites. We asked all states to verify the accuracy of their data, fill in missing information, and make corrections as needed.

- **Analyses were limited to years with comparable test data.** In recent years, many states have made changes to their testing systems that have created "breaks" in test data, meaning that test results after the change should not be compared with results from before the change. For example, test results should not be compared if a state has introduced a new test (and has not taken steps to equate the scores on the old and new tests). Data breaks can also occur if a state has changed the cut score for reaching a certain achievement level (such as proficient) on the test or has adopted a new scoring scale. To identify when a trend line had been broken, we gathered information from states about changes in their testing programs and excluded years for which data were not comparable.

- **The analysis focused on the period since 2002.** We limited our analyses to the period from 2002 (the year NCLB was signed into law) through 2008 (the most recent year for which most states had reported data at the time of the study). For many states, however, the trend lines begin later than 2002 because the state introduced new tests or made other changes that created a break in the comparability of test data. Because we had already reported on trends through 2006-07 in last year's report (CEP, 2008), states were included in the various analyses only if they had new data through school year 2007-08. Trends for 1999 through 2002, the years before NCLB was enacted, were available in only 13 states with comparable data and were discussed in CEP's first achievement report (CEP, 2007).

- **Trends were based on a minimum of three consecutive years of data.** Trends in test scores seldom move up or down in a straight line. More often they zigzag from year to year, even if the general trend is moving in one main direction (see CEP's 2007 report, pp. 16-17, for more about why test scores tend to fluctuate). Furthermore, tests are not perfect measures of achievement, so year-to-year variations in scores may be more attributable to measurement error than to actual changes in learning. Because of these factors, it can be misleading to base judgments about achievement on a single year of test results or to draw conclusions about the direction of a trend from the difference between two data points. For this reason, we considered a movement in test results to be a "trend" only if it was based on three or more years of consecutive years of data.

- **Trends were based on averages.** To further minimize any misleading effects of natural yearly fluctuations in test scores, and to make trends comparable whether they spanned three years or seven years or something in between, we averaged test data over three or more years. Because NCLB requires states to report test results in terms of the percent-

ages of students scoring at various achievement levels, the main indicator used in this study is the average annual percentage point gain or decline. This is the increase (or decrease) in the percentage of students scoring at a certain achievement level, divided by the number of years of testing minus one.

- **Changes in achievement were classified by size.** We distinguished between moderate-to-large gains and slight gains, using definitions developed for CEP's previous achievement studies with advice from the expert panel. An average gain or decline of 1.0 or more in the percentage of students scoring at a particular achievement level was considered a moderate-to-large change; an average gain or decline of less than 1.0 percentage point was considered a slight change. When calculations of average annual gains equaled 0.0, we considered this "no net change." For effect sizes, an average gain or decline of 0.02 standard deviation units or greater was considered a moderate-to-large change; less than 0.02 was considered slight; and 0.00 represented no change.

- **Grades for analysis were selected using consistent rules.** For time and complexity reasons we limited the analysis to three grades—one each at the elementary level (grade 4), the middle school level (usually grade 8), and the high school level (usually grade 10). If data were not available at any of these default grades, an adjacent grade with sufficient data was chosen according to a consistent set of rules.

- **An expert panel provided advice and technical assistance.** A panel of five nationally known experts in educational testing or education policy provided advice on all significant aspects of the study design, reviewed data, and commented on drafts of this report. The panel consisted of the following members:

  - 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
  - W. James Popham, professor emeritus, University of California, Los Angeles

Although the panel members, as well as HumRRO staff, provided input on this report, we did not ask them to endorse it, so the findings and views expressed here are those of CEP.

Appendix 2. State-by-State Trends at Three Achievement Levels

The following tables (A, B, and C) show the direction and size of test score trends in reading and math in each of the 50 states between 2002 (or a later year with comparable data) and 2008. **Table A** displays trends at the elementary grade analyzed, **table B** at the middle school grade, and **table C** at the high school grade. All three tables illustrate trends at three achievement levels: basic-and-above, proficient-and-above, and advanced. The tables use a set of symbols, listed in the legend below, to represent the size of the gains or declines.

LEGEND:

- **I = Moderate–to–large increase** *(≥1.0 percentage point average annual gain)*
- **i = Slight increase** *(<1.0 percentage point average annual gain)*
- **D = Moderate–to–large decrease** *(≥1.0 percentage point average annual decline)*
- **d = Slight decrease** *(<1.0 percentage point average annual decline)*
- **– = No net change** *(0.0 percentage point average annual change)*

Table A. ELEMENTARY SCHOOL: State–by–state trends on state tests

Percentages of students scoring at three achievement levels for years between 2002 and 2008 with comparable data

State	Elementary School Reading			Elementary School Math		
	Basic & Above	Proficient & Above	Advanced	Basic & Above	Proficient & Above	Advanced
AK	i	I	I	I	I	I
AL	–	i	i	i	I	I
AR	I	I	I	I	I	I
AZ	i	I	d	i	I	I
CA	I	I	I	i	I	I
CO	NA	i	d	NA	i	I
CT	D	D	i	i	i	I
DE	–	d	i	d	d	I
FL	I	I	I	I	I	I
GA	NA	I	I	NA	NA	NA
HI	NA	NA	NA	NA	NA	NA
IA	NA	i	D	NA	i	–
ID	NA	NA	NA	NA	NA	NA
IL	d	i	i	i	d	i
IN	NA	–	d	NA	–	i
KS	I	I	I	I	I	I
KY	NA	NA	NA	NA	NA	NA
LA	i	I	I	I	I	I
MA	d	d	i	I	I	I
MD	NA	I	I	NA	I	I
ME	i	I	–	I	i	–
MI	d	i	I	I	I	I

continued on page 25

State	Elementary School Reading			Elementary School Math		
	Basic & Above	Proficient & Above	Advanced	Basic & Above	Proficient & Above	Advanced
MN	D	D	D	i	i	I
MO	I	i	D	i	i	d
MS	NA	NA	NA	NA	NA	NA
MT	I	I	I	I	I	I
NC	NA	NA	NA	I	I	I
ND	d	i	I	d	d	d
NE	NA	I	NA	NA	I	NA
NH	I	I	I	I	I	I
NJ	NA	i	d	NA	I	I
NM	d	d	d	i	d	d
NV	I	I	d	I	I	I
NY	i	I	d	i	I	I
OH	i	I	d	D	D	D
OK	I	I	D	i	I	I
OR	NA	NA	NA	NA	NA	NA
PA	i	I	I	i	I	I
RI	I	I	I	I	I	–
SC	i	I	i	i	i	i
SD	–	I	d	–	i	i
TN	NA	I	I	NA	I	I
TX	NA	I	i	NA	I	i
UT	i	i	i	d	i	I
VA	NA	I	I	NA	I	I
VT	D	d	I	D	D	d
WA	d	I	I	d	i	i
WI	d	d	d	I	I	I
WV	i	I	i	i	I	I
WY	I	I	I	i	I	I
SUMMARY - Total number of states with increases, no net change, decreases, and insufficient data						
I	11	27	19	14	27	29
i	10	9	9	14	10	6
–	3	1	1	1	1	3
d	7	5	10	4	4	4
D	3	2	4	2	2	1
NA	16	6	7	15	6	7
Total	**50**	**50**	**50**	**50**	**50**	**50**

Table reads: Between 2002 and 2008, the percentage of Alaska students scoring at the basic level and above on the state elementary school reading test increased slightly. The percentage performing at the proficient level and above in elementary reading increased at a moderate-to-large rate, as did the percentage at the advanced level. On the state elementary school math test, moderate-to-large increases occurred in the percentages of Alaska students at all three achievement levels. Altogether, 11 states showed moderate-to-large increases in the percentage basic and above in elementary reading.

Note: Trends shown are for students in grade 4 in states with sufficient data at this grade. In other states, an adjacent elementary grade was selected according to a consistent set of criteria.

Table B. MIDDLE SCHOOL: State–by–state trends on state tests

Percentages of students scoring at three achievement levels for years between 2002 and 2008 with comparable data

State	Middle School Reading			Middle School Math		
	Basic & Above	Proficient & Above	Advanced	Basic & Above	Proficient & Above	Advanced
AK	i	I	I	I	I	I
AL	–	I	–	–	I	I
AR	I	I	I	I	I	I
AZ	–	I	i	i	–	–
CA	i	I	I	i	I	I
CO	NA	i	i	NA	I	I
CT	i	i	D	i	I	I
DE	d	D	D	I	I	I
FL	I	I	i	I	I	I
GA	NA	I	I	NA	NA	NA
HI	NA	NA	NA	NA	NA	NA
IA	NA	i	D	NA	i	I
ID	NA	NA	NA	NA	NA	NA
IL	d	I	d	i	I	i
IN	NA	i	d	NA	I	I
KS	I	I	I	I	I	I
KY	NA	NA	NA	NA	NA	NA
LA	i	I	d	I	I	i
MA	–	i	–	I	I	I
MD	NA	I	I	NA	I	I
ME	I	I	I	I	I	–
MI	I	I	I	I	I	I
MN	d	i	I	D	d	I
MO	I	I	d	I	I	i
MS	NA	NA	NA	NA	NA	NA
MT	I	I	I	d	D	I
NC	NA	NA	NA	I	I	I
ND	i	i	i	i	I	I
NE	NA	I	NA	NA	I	NA
NH	I	I	I	I	I	I
NJ	NA	NA	NA	NA	NA	NA
NM	I	I	i	I	I	I
NV	i	i	d	d	i	d
NY	I	I	i	I	I	I

continued on page 27

State	Middle School Reading			Middle School Math		
	Basic & Above	Proficient & Above	Advanced	Basic & Above	Proficient & Above	Advanced
OH	i	i	I	I	I	I
OK	I	I	i	I	I	I
OR	NA	NA	NA	NA	NA	NA
PA	I	I	I	I	I	I
RI	I	I	I	I	–	–
SC	i	i	i	I	i	–
SD	–	I	–	–	I	I
TN	NA	I	I	NA	I	I
TX	NA	I	I	NA	I	I
UT	i	I	I	I	I	I
VA	NA	I	I	NA	I	I
VT	–	I	I	d	d	i
WA	d	D	i	I	I	I
WI	–	d	I	i	i	i
WV	i	i	i	i	I	I
WY	I	I	D	I	I	I
SUMMARY - Total number of states with increases, no net change, decreases, and insufficient data						
I	14	29	20	22	34	32
i	10	11	10	7	4	5
–	6	0	3	2	2	4
d	4	1	5	3	2	1
D	0	2	4	1	1	0
NA	16	7	8	15	7	8
Total	**50**	**50**	**50**	**50**	**50**	**50**

Table reads: Between 2002 and 2008, the percentage of Alaska students scoring at the basic level and above on the state middle school reading test increased slightly. The percentage performing at the proficient level and above in middle school reading increased at a moderate-to-large rate, as did the percentage at the advanced level. On the state middle school math test, moderate-to-large increases occurred in the percentages of Alaska students at all three achievement levels. Altogether, 14 states showed moderate-to-large increases in the percentage basic and above in middle school reading.

Note: Trends shown are for students in grade 8 in states with sufficient data at this grade. In other states, an adjacent middle school grade was selected according to a consistent set of criteria.

Table C. HIGH SCHOOL: State–by–state trends on state tests

Percentages of students scoring at three achievement levels for years between 2002 and 2008 with comparable data

State	High School Reading			High School Math		
	Basic & Above	Proficient & Above	Advanced	Basic & Above	Proficient & Above	Advanced
AK	i	d	D	–	d	d
AL	–	D	D	–	I	i
AR	I	I	–	I	I	I
AZ	i	i	d	i	–	D
CA	I	i	d	I	I	I
CO	NA	i	i	NA	i	i
CT	NA	NA	NA	NA	NA	NA
DE	–	i	i	d	d	–
FL	d	i	I	I	I	I
GA	NA	NA	NA	NA	i	d
HI	NA	NA	NA	NA	NA	NA
IA	NA	i	d	NA	d	d
ID	NA	NA	NA	NA	NA	NA
IL	d	D	D	d	d	I
IN	NA	d	–	NA	i	d
KS	I	I	D	I	I	I
KY	NA	NA	NA	NA	NA	NA
LA	I	I	d	I	I	I
MA	I	I	i	I	I	I
MD	NA	NA	NA	NA	NA	NA
ME	d	I	i	NA	NA	NA
MI	NA	NA	NA	NA	NA	NA
MN	I	I	I	I	I	I
MO	D	D	i	I	I	i
MS	NA	NA	NA	NA	NA	NA
MT	I	I	I	d	D	–
NC	NA	NA	NA	NA	NA	NA
ND	D	D	D	i	i	D
NE	NA	I	NA	NA	I	NA
NH	NA	NA	NA	NA	NA	NA
NJ	NA	i	d	NA	I	i
NM	NA	NA	NA	NA	NA	NA
NV	D	d	I	D	D	I
NY	NA	NA	NA	NA	NA	NA

continued on page 29

State	High School Reading			High School Math		
	Basic & Above	Proficient & Above	Advanced	Basic & Above	Proficient & Above	Advanced
OH	I	I	d	I	I	I
OK	I	I	I	NA	NA	NA
OR	NA	NA	NA	NA	NA	NA
PA	i	I	I	i	I	i
RI	NA	NA	NA	NA	NA	NA
SC	i	i	d	i	i	i
SD	–	D	d	–	I	d
TN	NA	I	I	NA	i	d
TX	NA	I	I	NA	I	I
UT	i	I	i	i	I	I
VA	NA	I	I	NA	I	NA
VT	NA	NA	NA	NA	NA	NA
WA	I	I	i	i	I	i
WI	i	d	d	d	D	D
WV	d	d	d	d	I	i
WY	I	I	i	I	I	D
SUMMARY - Total number of states with increases, no net change, decreases, and insufficient data						
I	11	17	9	10	20	12
i	6	8	8	6	6	8
–	3	0	2	3	1	2
d	4	5	10	5	4	6
D	3	5	5	1	3	4
NA	23	15	16	25	16	18
Total	**50**	**50**	**50**	**50**	**50**	**50**

Table reads: Between 2002 and 2008, the percentage of Alaska students scoring at the basic level and above on the state high school reading (English language arts) test increased slightly. The percentage performing at the proficient level and above in high school reading decreased slightly, and the percentage at the advanced level decreased at a moderate-to-large rate. On the state high school math test, the percentage basic and above showed no net change, while the percentages scoring at the proficient level and above and the advanced level decreased slightly. Altogether, 11 states showed moderate-to-large increases in the percentage of students scoring at basic and above in high school reading.

Note: Trends shown are for students in grade 10 in states with sufficient data at this grade. In other states, an adjacent high school grade was selected according to a consistent set of criteria.

Credits and Acknowledgments

This report was written by Naomi Chudowsky, Victor Chudowsky, and Nancy Kober, CEP consultants. Naomi and Victor Chudowsky also analyzed the state test data for this study. A team from the Human Resources Research Organization (HumRRO) under the direction of Sunny Becker conducted extensive research and analysis for the report. 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, including the state profiles. Jennifer McMurrer, CEP's research associate, provided data analysis and quality control, oversaw the state profiles, and assisted with the project in general. Jack Jennings, CEP's president and CEO, provided advice and assistance. Other key HumRRO staff involved in the project were Hilary Campbell, Monica Gribben, and Wade Buckland.

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
1001 Connecticut Avenue, NW, Suite 522
Washington, D.C. 20036
tel: 202.822.8065
fax: 202.822.6008
e: cep-dc@cep-dc.org
w: www.cep-dc.org