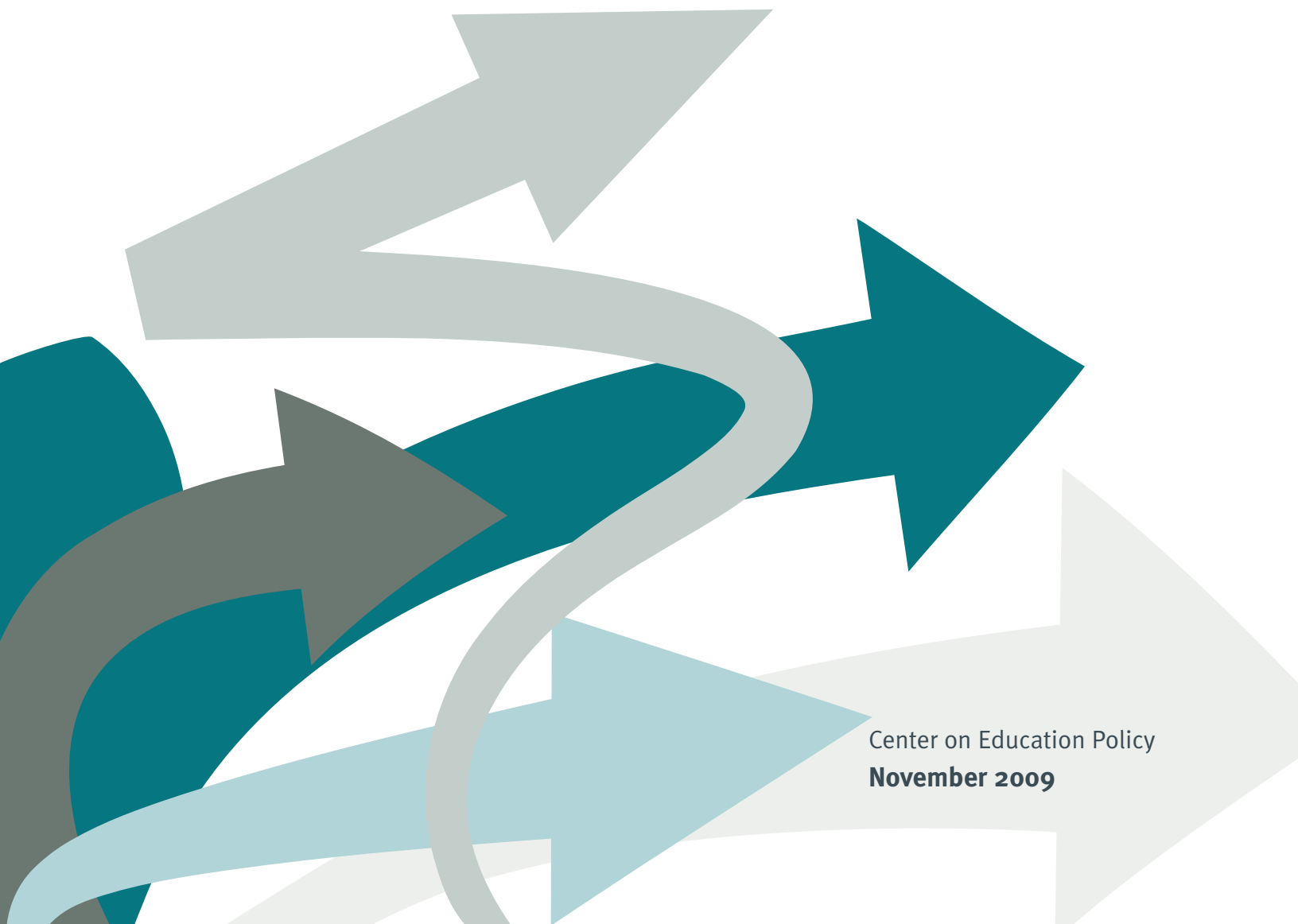


State Test Score Trends Through 2007-08, Part 4

Has Progress Been Made in Raising Achievement for Students with Disabilities?



Center on Education Policy
November 2009

State Test Score Trends Through 2007-08, Part 4

Has Progress Been Made in Raising Achievement for Students with Disabilities?

Table of Contents

Introduction	1
Main Findings	1
Background and Study Methods	2
Problems in Discerning Achievement Trends for Students with Disabilities.....	4
Grade 4 Trends for Students with Disabilities at Three Achievement Levels	10
Differences between Students with Disabilities and Other Students at the Proficient Level	13
Conclusion	14
References	16

List of Tables

Table 1.	Percentage of students with disabilities who took and scored proficient on various types of assessments in elementary reading, 2007	6
Table 2.	Percentage change in the number of students with disabilities taking state tests between 2006 and 2008	8
Table 3.	Number (and percentage) of states showing various trends for students with disabilities at three achievement levels on state grade 4 tests	11
Table 4.	Percentages of students with disabilities and all tested students in grade 4 reaching various achievement levels, 2008	12
Table 5.	Percentages proficient for students with disabilities and non-disabled students at three grade levels, 2008	13

Has Progress Been Made in Raising Achievement for Students with Disabilities?

Introduction

The federal No Child Left Behind Act (NCLB) holds schools, districts, and states accountable for improving the academic achievement of all students, including the nearly 14% of public school students who receive special education services because they have an identified disability. By 2014, 100% of students with disabilities are expected to perform at the “proficient” level on state tests—the same goal set for students overall and for other subgroups of students. NCLB also calls on schools and districts to close achievement gaps between students with disabilities and their non-disabled peers and to include these students in regular state testing programs to the maximum extent possible.

This report by the Center on Education Policy (CEP) examines progress in raising achievement for students with disabilities. It also describes the factors that make it difficult to clearly discern achievement trends for this subgroup. The data for this analysis were collected by CEP with technical support from the Human Resources Research Organization, and come from the state reading and mathematics tests used for NCLB accountability in all 50 states. The trends discussed in this report cover fewer years—2006 through 2008—than the trends in the three earlier parts of this 2009 series of reports on student achievement. This is because federal policies for testing students with disabilities and determining their progress were revised in 2003 and 2005 in ways that likely affected the year-to-year comparability of test data. School year 2005-06 marked the first year that states administered tests under the revised rules, and so we now have enough data to establish a three-year trend line for students with disabilities in many states.

Main Findings

Several main findings emerged from our study:

- **Fuzzy data make it difficult to obtain a clear picture of achievement for students with disabilities.** States administer two or three types of assessments to these students—the regular state test (with or without test accommodations) and one or two types of alternate assessments—each with its own definition of proficient performance. Many states have yet to demonstrate that alternate assessments for this group are reliable and that interpretations of their results are valid. In addition, the percentage of students with disabilities tested with alternate assessments varies widely from state to state and year to year, and states differ as to whether and how they report the results. Even on regular state tests, accommodations allowed for students with disabilities might affect results. Furthermore, the number of students tested in this subgroup has changed rapidly since 2006 in many states, which could affect achievement trends if the new students in the group have different characteristics. As a result of these and other factors, test data for students with disabilities are less precise than one would like, but the data still provide a rough indicator of achievement trends for this group.

- **Students with disabilities 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 this subgroup. In reading, for example, the percentage of students with disabilities scoring at the advanced level increased in 25 of the 41 states with sufficient data and decreased in 11 states. In math, 26 of 42 states showed gains at the advanced level while 11 showed a decline. Overall, the proportion of states with gains for students with disabilities at the three achievement levels was roughly similar to the proportions for all students and for students in the racial-ethnic and low-income subgroups tracked for NCLB.
- **Differences in test performance between students with disabilities and non-disabled students are very large.** Differences in the percentage proficient between these groups often exceed 30 or even 40 percentage points in reading and math.

Readers who are interested in trends for students with disabilities in a specific state are encouraged to access the detailed state profiles of subgroup achievement trends available on CEP's Web site (www.cep-dc.org).

Background and Study Methods

About 13.6 % of public school students in grades prekindergarten through 12 receive federal education services for students with disabilities, according to recent data from the U.S. Department of Education (ED) (2009). This group encompasses learners who have a wide range of academic aptitudes and whose disabilities differ in both type and severity. About 40% have specific learning disabilities, about 22% have speech or language impairments, about 8% have mental retardation, and the rest have visual or hearing impairments, emotional disturbances, autism, or other disabilities.

People often assume that most of these students, by definition, are unable to perform at the proficient level on state tests. This is not the case. By some estimates, between 75% and 80% of students classified as having disabilities do not have intellectual impairments.¹ With appropriate accommodations—such as Braille or large-print tests, individual or small group test administrations, and more time to complete a test—many students with disabilities can perform at or above the proficient level on tests geared to their grade level. At the same time, a subset of these students struggles with taking regular state tests and achieving proficient scores, even with accommodations.

Since 2002, educators and policymakers have wrestled with issues of how to meet the NCLB achievement goals for students with disabilities—particularly those whose disabilities are significant—while implementing the law's testing and accountability requirements in a fair and appropriate way. Policy decisions related to this issue are further complicated by the philosophical conflict that exists between the goals of NCLB and the Individuals with Disabilities Education Act (IDEA), the federal law that governs how states provide special education services. Recognizing the unique and varied needs of students with disabilities, IDEA requires schools to develop an individualized education plan (IEP) that defines educational

¹ Martha Thurlow, director of the National Center on Educational Outcomes, estimates that about 80% of students with disabilities—those with learning disabilities, speech or language impairments, emotional or behavioral difficulties, and visual, hearing and other physical impairments—do not have the types of intellectual impairments that would prevent them from demonstrating proficiency at grade level, as long as they had appropriate accommodations and instruction. Somewhat less than 20% of students with disabilities have intellectual impairments, including students with mental retardation, some with multiple disabilities, and some with autism. Even many of these students can achieve proficiency. See Quenemoen, 2009, p. 159.

goals for each participating student and spells out services to be provided to meet those goals. Schools must then measure a student's progress against his or her individual goals. NCLB, however, measures that same student's progress relative to goals established for all students and requires most students with disabilities to take the same state reading and math tests as other students in their grade. In short, NCLB implies a uniform approach to assessment and educational goals, while IDEA implies an individualized one.

In the first few years after the law's passage, making adequate yearly progress (AYP) for students with disabilities was a major issue in NCLB implementation, according to CEP surveys of state and district officials (CEP, 2004; 2005). As explained in more detail in the next section, the U.S. Department of Education made policy changes that allowed a limited number of students with disabilities to be tested using various types of alternate assessments instead of the regular state test.² Although these changes brought some welcome flexibility to testing and accountability requirements for this subgroup, states and school districts continue to face challenges in designing and implementing alternate assessments for these students and helping this subgroup, which is often the lowest-performing in a state, meet achievement targets. At the same time, advocates for students with disabilities have urged Congress to "stay the course" and maintain key aspects of the law that hold school systems accountable for the performance of these students and encourage their inclusion in regular assessments and instructional programs (Thurlow, 2004; Quenemoen, 2006; National Disability Rights Network, 2007; Advocacy Institute, 2007; Consortium for Citizens with Disabilities, 2007).

While problems associated with making AYP and testing students with disabilities under the law are well known, how this group is faring in terms of achievement is less well understood. This report, the fourth in a 2009 series of CEP reports on student achievement trends,³ is our first to focus specifically on the achievement of students with disabilities, although we have covered issues relating to these students since our first report on NCLB implementation in 2003.

The National Center on Educational Outcomes (NCEO) at the University of Minnesota has also been monitoring the performance of students with disabilities on state tests for much of the past decade. Each year, NCEO presents a snapshot of the percentages of these students scoring at the proficient level in a given year, along with breakdowns of the numbers of these students taking regular or alternate assessments. One NCEO study, a 2005 survey of state special education coordinators, found that most state officials were seeing gains in achievement for this subgroup. The respondents attributed this increase to, among other things, better alignment of IEPs to state standards, improved data collection, increased access to standards-based instruction, greater participation in assessments, requirements for these students to achieve at their grade level, and the desire to avoid NCLB sanctions (Thompson et al., 2005).

² As specified by the U.S. Department of Education in the preamble to regulations published on December 9, 2003, an alternate assessment "may include materials collected under several circumstances, including (1) teacher observation of the student, (2) samples of student work produced during regular classroom instruction that demonstrate mastery of specific instructional strategies in place of performance on a computer-scored multiple-choice test covering the same content and skills, or (3) standardized performance tasks produced in an 'on-demand' setting, such as completion of an assigned task on test day" (U.S. Department of Education, 2003). These regulations emphasize that these are just examples, however. Alternate assessments may also take the form of modified or simplified paper-and-pencil tests, and states have considerable flexibility in designing the most appropriate format.

³ Part 1 in the series of reports, *State Test Score Trends Through 2007-08*, examined trends since 2002 at the basic, proficient, and advanced levels of achievement for students as a whole and found that many more states had gains than declines at all three achievement levels (CEP, 2009a). Part 2 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). Part 3 found that achievement gaps on state tests between different racial/ethnic groups of students and between low-income and more advantaged students have tended to narrow since 2002 (CEP, 2009c). These reports can be downloaded from CEP's Web site at www.cep-dc.org.

Our study differs from those of NCEO in that we examined test score trends spanning three consecutive years to see whether achievement for students with disabilities moved upward or downward between 2006 and 2008, the latest year for which test results were available at the time we collected our data. Our main unit of analysis is the trend line, which is a record of change from 2006 through 2008 in the performance of the subgroup of students with disabilities at a particular achievement level in one subject and grade level in one state. For example, the change during this period in the percentage proficient in reading for 4th graders with a disability in California constitutes one trend line. This report includes only trends based on three consecutive years of comparable test data—the minimum span needed to establish a direction in achievement. States have been omitted if they made major changes in their testing system after 2005-06, such as switching tests or changing cut scores for various achievement levels, or if they lacked data for other reasons.

To determine whether students with disabilities have made progress across the achievement spectrum, we looked at grade 4 trends for this subgroup at the basic-and-above, proficient-and-above, and advanced achievement levels. In particular, we calculated the average annual percentage point gain or decline for each trend line and counted the number of states showing gains or declines at each achievement level in reading and math. (The average annual gain or decline is simply the increase or decrease in the percentage of students scoring at or above a certain achievement level divided by two, because we are looking at the degree of change between school years 2005-06 and 2007-08.) We limited this analysis to one elementary grade because of the massive amount of data involved and because this was the pilot year for a process that CEP hopes to expand to the middle and high school levels in future years.

To provide a rough idea of how the subgroup with disabilities is performing relative to other students, we also summarized data on the percentages of students with disabilities and non-disabled students reaching the proficient level in each state in 2008 at three grade levels: elementary (grade 4 in all cases), middle school (usually grade 8), and high school (generally grades 10 or 11).

We conducted this study with advice from a panel of five nationally known experts in educational testing or education policy.⁴ More details about study methods can be found in appendix 1 to part 1 of this series of reports (CEP, 2009a).

Problems in Discerning Achievement Trends for Students with Disabilities

Finding: State test data for students with disabilities are sometimes fuzzy, which makes it difficult to get a clear picture of achievement trends for this group.

Achievement data for students with disabilities can be somewhat imprecise for several reasons. The most notable ones are 1) changes in federal policies for testing and determining AYP for students with disabilities and variations in states' implementation of these changes; 2) questions about whether regular assessments with accommodations and various alternate

⁴ 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.

assessments meet standards for reliability and yield valid interpretations of results; and 3) rapid shifts in the numbers of test-takers with disabilities that could affect the comparability of test results in the same state.

ADEQUATE YEARLY PROGRESS DETERMINATIONS AND ALTERNATE ASSESSMENTS

Not long after NCLB was enacted, many educators and state and local officials began to question the educational value and fairness of assessing all students with disabilities with regular tests, particularly students with significant cognitive disabilities. In school year 2003-04, 37% of schools nationwide with subgroups of students with disabilities large enough to count⁵ did not make AYP for this subgroup (Stullich et al., 2007). Thirteen percent of schools that did not make AYP that year fell short *solely* because of the test performance of the subgroup with disabilities. States protested because they felt these numbers were sure to escalate in future years as states' AYP targets increased.

To address the conflict between the emphasis on uniform standards and tests in NCLB and the need for individualization in IDEA, the U.S. Department of Education made policy changes in 2003 and 2005 that allowed for two exceptions to the use of regular state tests for students with disabilities. States, in turn, implemented these policies in various ways.

The first exception affected students with the most significant cognitive disabilities. States must still test these students, but they can use alternate assessments geared to students' learning level ("alternate standards") rather than their grade level. When states calculate the percentage of students scoring proficient on state tests for AYP purposes, no more than 1% of proficient scores can come from alternate assessments aligned to alternate standards (U. S. Department of Education, 2003). In certain circumstances, the Secretary of Education may allow a state—and a state may allow a school district—to exceed this 1% cap.

For many states, the 1% rule was not a sufficient solution, so ED devised a second exception for students with disabilities, an additional category of alternate assessments aligned to "modified standards"—described in federal guidance as standards that are aligned to content standards for the tested student's grade level and are "challenging" but may be "less difficult than grade-level achievement standards" (U.S. Department of Education, 2007). Tests aligned to modified standards often use test items with a simpler format than items in the regular state tests (for examples from a modified standard assessment in Kansas, see CEP, 2005). This policy change was intended to address the problem of "gap kids"—students whose cognitive disabilities are not severe but who nevertheless are not ready to take the regular state tests for their grade level. ED allowed proficient scores from this group to account for no more than an additional 2% of the tested population. However, according to a state survey conducted by the Government Accountability Office (GAO), only eight states administered alternate assessments aligned to modified standards in the 2007-08 school year (GAO, 2009).

In short, the proficient scores that states report for students with disabilities may come from as many as three different assessment systems—the regular test (with and without accommodations) and two types of alternate tests, one aligned to "alternate" and another to "modified" standards. For analyses of achievement trends, this represents something of a problem because each of these assessments is likely to have a different definition of "proficient" performance. States are not required to publicly break down test results for students with disabilities by the

⁵ Under NCLB, states may set minimum sizes defining how many students must be in a subgroup in order for that subgroup's performance to count separately in AYP determinations.

type of test taken. Furthermore, the information we collected for this study and our conversations with officials in a limited number of states suggest that some states do not include *any* scores from alternate assessments in the test data they report for NCLB purposes.

NCEO found that the types of tests taken by students with disabilities who achieved proficient scores varied widely among states. **Table 1** shows which assessments were the sources of proficient scores for students with disabilities in 2007 in four states with diverse approaches.

Table 1. Percentage of students with disabilities who took and scored proficient on various types of assessments in elementary reading, 2007

State	Total percentage of SWDs scoring proficient	Proficient scores from regular assessment	Proficient scores from alternate assessment, alternate standards	Proficient scores from alternate assessment, modified standards
Hawaii	11.3%	9.2%	2.1%	0.0%
Kansas	72.9%	47.9%	6.3%	18.7%
Nebraska	78.2%	74.4%	3.7%	0.0%
Texas	69.9%	23.7%	20.5%	25.7%

Table reads: In 2007, 11.3% of test-takers with disabilities in Hawaii scored at the proficient level in grade 4 reading on one of two types of exams: 9.2% of test-takers with disabilities took and scored proficient on the regular state assessment, and 2.1% took and scored proficient on an alternate assessment aligned to alternate standards. Hawaii did not offer an alternate assessment aligned to modified standards in 2007, so no students took advantage of this third type of exam.

Source: Thurlow, Altman, & Vang, 2009.

A closer look at Kansas and Nebraska, two neighboring states shown in table 1, illustrates how the percentage proficient for this subgroup may be derived from a very different mix of assessments. In both states, roughly three-fourths of test-takers with disabilities scored at the proficient level (72.9% in Kansas and 78.2% in Nebraska). In Kansas, one of the first states to develop an alternate assessment aligned to modified standards, roughly one-fifth (18.7%) of the tested population of students with disabilities took and scored proficient on the state's alternate assessment geared to *modified* standards, while 6.3% took and scored proficient on the state's alternate assessment geared to *alternate* standards. Less than half (47.9%) of the tested students with disabilities in Kansas took and achieved a proficient score on the regular state test. Nebraska, like the majority of states, has not developed an alternate assessment aligned to modified standards, so none of its students with disabilities could use that option. About 3.7% of Nebraska's test-takers with disabilities sat for and scored proficient on the assessment aligned to alternate standards. The regular state test was by far the main source of proficient scores for the subgroup with disabilities; three-fourths of these students took it and reached proficiency.

VALID AND RELIABLE ASSESSMENTS

As the GAO (2009) survey made clear, Nebraska is not alone in its lack of an alternate assessment aligned to modified standards. Not only are alternate assessments expensive to design and implement in general, but states have also had difficulty demonstrating that their alternate assessments for students with disabilities are sufficiently reliable and yield results that can be interpreted in a valid way. It is not clear to what extent these validity and reliability problems affect the data reported for NCLB, but they almost surely affect it in some way.

As of January 2009, 12 states' assessment systems had not met the expectations for ED's peer review process, and problems with alternate assessments for students with disabilities were an issue in 11 of these cases (GAO, 2009). GAO identified several issues with assessments aligned to alternate standards. First, it is not easy to develop a single test for a group of students with such a wide range of disabilities; a certain degree of individualization is necessary, which may affect validity and reliability. Second, the tests are administered to a relatively small population of students, so states have trouble gathering evidence of their validity and reliability. Finally, alternate assessments are more costly to develop and administer than regular state assessments. In Maryland, for example, it costs \$30 per pupil to administer the regular state test, but it costs from \$300 to \$400 per pupil to administer alternate assessments.

Another set of issues that affects the interpretation of test results for students with disabilities pertains to the use of accommodations on regular state tests. Koretz and Barton (2003) have noted a number of potential problems with accommodations, which are meant to improve the validity of score interpretations for individual students by acting as a kind of "corrective lens" that adjusts for a student's disabilities. Federal and state regulations demand that students with disabilities be tested with "appropriate" accommodations. But there is little agreement among researchers about how accommodations affect the validity of inferences drawn from test results and which types of accommodations are appropriate for which students. In the case of students who are visually impaired, the choices of accommodations are fairly straightforward and are likely to include large print versions of tests, audio recordings, or Braille versions. In other cases, however, such as accommodations for students with emotional disturbances or speech and language impairments, the match between the type of disability and accommodation is much less straightforward. Furthermore, debate persists about how certain commonly used accommodations, such as giving students more time to complete a test or reading questions aloud, affect scores and whether they can alter the type of knowledge or skill that is being measured.

Decisions about which students are identified for special education, what services are spelled out in their IEPs, and how they are to be included in the testing regimen are all made locally, so a great deal of variation can exist among schools, districts, and states. Although tests are standardized, accommodations are not. In an accountability system that emphasizes inclusion of students with disabilities in regular instruction and assessments and continuous increases in test scores, administrators might be tempted to overuse accommodations to include more students, or might offer accommodations to students who may not have received them in the past, thus biasing results. In addition, if accommodation practices change over time so that more students are included in a testing program, this makes it quite problematic to make comparisons over multiple years in the performance of the subgroup of students with disabilities. This has been an issue with the National Assessment of Educational Progress (NAEP), for example (Kitmitto & de Mello, 2008).

CHANGING NUMBERS OF TEST-TAKERS

The process of determining trends over time in achievement for students with disabilities is further complicated by rapid changes in the number of test-takers in this subgroup, according to data collected for our study. Between 2006 and 2008, some states have experienced disproportionately large fluctuations, either up or down, in the number of test-takers with disabilities, while other states have experienced small changes that were more in line with shifts in the overall tested population.

Table 2 gives an overview of changes in the number of test-takers with disabilities. For each subject and grade level, the table shows the largest percentage increase and decrease in this population found in any state, the median change in this population across all states with data (meaning that half the states were below this median figure and the other half above), and the number of states with increases or decreases of more than 20%.

Table 2. Percentage change in the number of students with disabilities taking state tests between 2006 and 2008						
Statistic	Elementary reading	Elementary math	Middle reading	Middle math	High school reading	High school math
Largest increase in any state	90%	107%	107%	89%	64%	85%
Largest decrease in any state	-32%	-22%	-26%	-30%	-23%	-34%
Median for all states with data	1%	1%	-3%	-3%	0%	2%
# of states with increase >20%	4	4	3	5	3	5
# of states with decrease >20%	2	2	3	5	1	1
# of states with data	47	46	47	46	40	40

Table reads: Between 2006 and 2008, the number of students with disabilities taking tests in elementary reading increased by as much as 90% in one state (the largest increase) and decreased by as much as 32% in another state (the largest decrease); the median change in the tested population with disabilities across the 47 states with data was a 1% increase. Four states experienced increases of more than 20%, and two states experienced decreases of more than 20% in the tested population of students with disabilities.

Source: Center on Education Policy based on data collected from state departments of education.

Altogether, we found 12 states in which the population of tested students with disabilities increased or decreased by more than 20% at one or more grade level/subject combinations; these states included Arizona, California, Florida, Kansas, Louisiana, Missouri, North Carolina, Nevada, Tennessee, Texas, Vermont, and Washington. In Texas, the number of test-takers with disabilities roughly doubled in this short period at some grade levels, while the overall population of test-takers increased by just 7% to 10%. We found many additional instances of changes in the number of test-takers with disabilities that were disproportionate to the change in the overall tested population. In elementary reading, for example, 14 of the 47 states with data showed changes in the number of test-takers with disabilities that did not closely match the extent of change in the overall tested population.

In Michigan, to cite one instance, the overall number of 4th graders tested in reading declined by 2.2%, but the number of tested students with disabilities increased by 10.8%. All told, the tested population of students with disabilities is a rapidly shifting one, which creates uncertainty about whether changes in test scores over three years are due to differences in the composition of the subgroup rather than changes in learning.

We contacted education officials in several states with particularly large increases and decreases to find out why these shifts in numbers of test-takers with disabilities were occurring. Their responses varied, but most were in some way related to shifts in the types of tests administered to students with disabilities, often at the behest of ED. The tested population of students with disabilities should be seen as the number who took state tests *considered valid by ED*, which in many cases is not the same as the actual number of students with disabilities who took any state test. For example, Vermont was administering two alternate assessments to students with disabilities but dropped one of them because ED disallowed the results; more students with disabilities were shifted over to the regular state test, and their scores were counted. Texas introduced its alternate assessment aligned to modified standards in spring of 2008, in keeping with ED's 2% rule, which meant that more students with disabilities were tested.

In Arizona, several factors affected the shift in the tested population. First, large increases occurred in the tested population of students with disabilities in middle and high school math after the state stopped allowing students with disabilities at these grade levels to use calculators on math tests. Calculators were considered a non-standard accommodation, and ED would not count the scores of students tested with this accommodation. When Arizona ended that practice, more scores from students with disabilities were counted between 2006 and 2008. Second, the state made an effort, through mandatory training of district test coordinators, to ensure all students with disabilities were tested, especially at the high school level (NCLB requires at least 95% of the students in each subgroup to participate in testing). Third, overall enrollment in Arizona's schools increased by 20% between 2006 and 2008, and some of this increase may have been reflected in the tested population of students with disabilities.

THE RESULT: A FUZZY PICTURE OF ACHIEVEMENT FOR STUDENTS WITH DISABILITIES

The No Child Left Behind Act gave states the latitude to define what constitutes proficient performance, as well as performance at the basic and advanced levels, in terms of their own tests, cut scores, and content standards.⁶ The result is 50 different definitions of proficiency in math and reading. But for students with disabilities, there are at least 108 definitions of proficiency at each grade level—one definition for the regular test and a second one for the alternate assessment aligned to alternate standards in 50 states, plus a third definition for the alternate assessment aligned to modified standards in another 8 states. For this study, we have published the percentage of students with disabilities who were *publicly reported and verified by states* to have scored proficient on state tests, regardless of the type of test they took. Therefore, the reader should bear in mind that although most students with disabilities are tested using regular state assessments, often they are using a variety of accommodations. In addition, varying percentages of proficient scores come from alternate assessments. In most states, the percentage of students with disabilities who took and scored proficient

⁶ In addition, some states use different names for the three achievement levels. More information about the process and challenges of analyzing performance by achievement level, as well as a detailed discussion of trends at three achievement 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).

on alternate assessments, aligned to either alternate or modified standards, ranged from 2% to 10%, but there were outliers, such as 46% in Texas, 25% in Kansas, and 20% in Oklahoma (NCEO, 2009). All three of these states have alternate assessments aligned to modified standards.

Because of the differences among states in approaches for testing students with disabilities and reporting results, comparisons between specific states should not be made. In addition, all of the complicating factors described above can lend an unknown degree of fuzziness even within the same state. Moreover, because ED was still refining testing policies for students with disabilities in the early years of NCLB, the trends in this report cover just three years, the minimum span used to determine trends for our achievement studies and a rather short span on which to base conclusions. Consequently, the trends discussed in this report are the best that can be gleaned from the data available, but one should keep in mind they are estimates rather than precise measures.

Grade 4 Trends for Students with Disabilities at Three Achievement Levels

***Finding:* The subgroup of students with disabilities has 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 these students at all 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.) Although the percentage of students scoring at or above the proficient level is the statistic used to determine AYP, the percentages scoring at the basic and advanced levels can also reveal valuable information about the performance of students with disabilities. For example, if the percentage of students with disabilities who reach or exceed the basic level in math has increased over time, this indicates that more students in this group have improved enough to move from the below-basic into the basic category—progress that would not show up if one looked only at the percentage proficient. Likewise, if more students with disabilities have moved from the proficient to the advanced category, this improvement would not be evident from the percentage proficient alone.

To better understand how students with disabilities are performing across the achievement spectrum, we looked at the number of states with gains or declines from 2006 to 2008 in the percentages of 4th graders with disabilities scoring at the basic-and-above, proficient-and-above, or advanced levels on state tests. Under the approach we used, 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 trends that can become quite complex.

Table 3 presents the results of our analysis of trends by achievement level in reading and math. For example, 41 states had adequate data to discern a trend between 2006 and 2008 in the percentage of students with disabilities scoring at the advanced level in reading. Twenty-five of these states showed gains for this subgroup at the advanced level, and 11 states showed declines. The percentages of states with gains did not vary greatly across the three achievement levels; however, the number of states with gains was greatest at the proficient level, largely because more states had sufficient data to determine a trend at this level. Overall, the ratios of gains to declines for students with disabilities at the three achievement levels are roughly similar to those we found for other subgroups of students, as described in detail in part 3 of this series of reports (CEP 2009c).⁷

Table 3. Number (and percentage) of states showing various trends for students with disabilities at three achievement levels on state grade 4 tests

Achievement level and trend	Reading	Math
Basic-and-above		
Gain	19 (58%)	24 (69%)
Decline	11 (33%)	11 (31%)
No change	3 (9%)	0 (0%)
Total # of states with data	33	35
Proficient-and-above		
Gain	28 (65%)	29 (66%)
Decline	11 (26%)	15 (34%)
No change	4 (9%)	0 (0%)
Total # of states with data	43	44
Advanced		
Gain	25 (61%)	26 (62%)
Decline	11 (27%)	11 (26%)
No change	5 (12%)	5 (12%)
Total # of states with data	41	42

Table reads: Of the 33 states with sufficient data to analyze trends in grade 4 reading, 19 states, or 58%, showed gains in the percentage of students with disabilities scoring at or above the basic level of achievement; 11 states, or 33%, showed declines at this level; and 3 states, or 9%, showed no net change.

Source: Center on Education Policy based on data collected from state departments of education.

⁷ Other subgroups that are the focus of NCLB may also include some students with disabilities. For example, a Latino student with a disability who comes from an economically disadvantaged background would be included in the Latino and low-income subgroups, as well as in the subgroup of students with disabilities.

Table 4 gives a snapshot across states of the percentages of 4th graders with disabilities performing at the three achievement levels in 2008, the most recent year of test data collected for this report. As the table illustrates, the median, or middle, percentage of students with disabilities reaching the basic-and-above level of achievement exceeded 70% in both reading and math, but the median reaching the proficient-and-above level was under 50% in both subjects. The median reaching the advanced level was 8% in reading and 10% in math.

Table 4. Percentages of students with disabilities and all tested students in grade 4 reaching various achievement levels, 2008						
Statistic	Basic & above		Proficient & above		Advanced	
	SWDs	All students	SWDs	All students	SWDs	All students
Reading						
Median	73%	92%	42%	74%	8%	26%
Lowest % in any state	31%	79%	13%	46%	0%	4%
Highest % in any state	100%	100%	79%	92%	33%	60%
# of states with sufficient data	33	34	43	44	41	43
Math						
Median	71%	90%	49%	75%	10%	28%
Lowest % in any state	32%	76%	10%	39%	2%	7%
Highest % in any state	100%	100%	85%	94%	34%	59%
# of states with sufficient data	35	35	44	44	42	42

Table reads: For the 33 states with sufficient data, the median percentage of students with disabilities achieving at the basic level or above in grade 4 reading was 73%; the comparable median for all students reaching the basic level or above was 92%. The lowest percentage of students with disabilities scoring at basic-and-above level in grade 4 reading found in any state was 31%; the highest percentage found in any state was 100%.

Source: Center on Education Policy based on data collected from state departments of education.

Table 4 also shows the spread across states in the percentage of students with disabilities attaining each of these three achievement levels. In one state no students with disabilities scored at the advanced level in reading, while in another state 33% did so. These disparities are often a product of different state policies for students with disabilities and state testing systems that vary greatly in content, rigor, and other aspects, so this range should not be viewed as an indicator that students in this subgroup are learning much more in some states than in others.

Additionally, table 4 displays results at three achievement levels for *all* tested students—a group that also includes students with disabilities. However, a better way to gauge how students with disabilities are performing relative to other students can be found in table 5 in the next section, which looks at differences between students with disabilities and non-disabled students.

We also analyzed trends for the middle and high school grades at the *proficient level only*. As with the elementary level, more states showed gains than declines in the percentage proficient for students with disabilities at the middle and high school grades analyzed.

Differences between Students with Disabilities and Other Students at the Proficient Level

Finding: Differences in state test performance between students with disabilities and non-disabled students remain large.

In part 2 of this series of reports, we reported on progress in narrowing gaps for racial/ethnic minority students and low-income students, and we had initially planned to do a similar analysis of gap trends for students with disabilities. But we concluded that the problems of fuzzy data for this subgroup would be magnified if we tried to compare *trends over time* between two groups subject to different testing procedures—students with disabilities, who take one of three types of tests and may use accommodations, and non-disabled students, who take the same regular test and typically do not use accommodations. We did not have faith that such a comparison would lead to accurate or valid findings about changes in the size of the gap.

We do think it is important, however, to give some indication of how students with disabilities are doing relative to other students. Rather than analyzing gaps over time, we decided that a more careful approach would be to give a snapshot from one year. **Table 5** shows the percentages of students with disabilities and non-disabled students at three grade levels who scored at the proficient level on state tests in 2008. The percentage proficient is the most important determinant of adequate yearly progress under NCLB.

Table 5. Percentages proficient for students with disabilities and non-disabled students at three grade levels, 2008

Statistic	Elementary		Middle school		High school	
	SWDs	Non-disabled	SWDs	Non-disabled	SWDs	Non-disabled
Reading						
Median	41%	79%	34%	78%	31%	77%
Lowest % in any state	14%	49%	3%	30%	4%	41%
Highest % in any state	79%	95%	78%	96%	85%	98%
# states with sufficient data	43		43		34	
Math						
Median	49%	79%	28%	74%	22%	69%
Lowest % in any state	16%	42%	3%	23%	5%	37%
Highest % in any state	85%	94%	72%	94%	69%	96%
# states with sufficient data	44		43		33	

Table reads: For the 43 states with sufficient data, the median percentage of students with disabilities scoring at the proficient level or above in grade 4 reading was 41% in 2008, compared with a median of 79% for students who did not have disabilities. The lowest percentage of students with disabilities scoring proficient in grade 4 reading in any state was 14%; the highest was 79%.

Note: The figures in table 5 differ from those in table 4 because the comparison groups are different—students who are not disabled in the case of table 5, and all tested students (including students with disabilities) in table 4. In addition, table 4 includes only elementary school students.

Source: Center on Education Policy based on data collected from state departments of education.

As the table reveals, the differences in percentages proficient between students with and without disabilities are quite large—disparities of 30 or even 40 percentage points are common. Of the 43 states with sufficient data in elementary reading, 28 states had differences of more than 30 percentage points between these two groups. The median rows in table 5 illustrate the magnitude of this difference from another perspective. In middle school math, for example, the median percentage proficient for students with disabilities was 28% in 2008, signifying that half the states were below 28% and half were above. The median for non-disabled students was 74%—a difference of 46 percentage points. In high school reading, the respective medians were 32% and 75%, a difference of 43 points. Taken together, these figures show the difficulty many states face in getting this subgroup to 100% proficiency and suggest why ED and the states made adjustments and allowances in NCLB testing rules.

Table 5 also shows the remarkable range among states in their percentages of students with disabilities reaching proficiency in various subjects and grade levels. The spread between the lowest and highest percentage proficient in any state was greater for students with disabilities than for non-disabled students. In high school reading, for example, just 4% of students with disabilities in one state scored high enough on the state test to be considered proficient in 2008, while in a neighboring state 85% of students with disabilities were considered proficient. These differences are partly due to the fact that states have different standards, assessments, and cut scores. But in the case of students with disabilities, the even larger disparities between the minimum and maximum states may be attributable to different policies about which students are included in this subgroup, which assessments they take, and how alternate assessments are scored.

Conclusion

The trends described in this report suggest that students with disabilities are making progress on state tests. Average test scores for this subgroup have increased, and the percentages of students reaching the basic, proficient, and advanced levels have risen in most states. Still, the differences in state test performance between students with disabilities and non-disabled students remain very large. At the same time, we must reiterate that the findings in this report are a rough estimate of achievement trends for students with disabilities rather than a precise description. That is because the numbers of students with disabilities scoring proficient reported by states are fuzzy for reasons described above.

Complexities with testing policies and achievement data for students with disabilities are not limited to state tests. Recently, the National Assessment Governing Board (NAGB), which oversees the federal National Assessment of Educational Progress, grappled with how to increase and standardize the inclusion of students with disabilities in NAEP testing. As with NCLB and state testing, NAGB looked at ways to include maximum numbers of students with disabilities in NAEP while recognizing that some students within this group do have significant disabilities. A technical advisory panel recommended that NAGB first ensure maximum participation—states should include in NAEP at least 95% of the students with disabilities drawn for the NAEP sample. Second, students with disabilities who qualify to take alternate assessments aligned to alternate standards for state testing should be excluded from NAEP, with a target of excluding not more than 1% of the sample. Those students who take alternate assessments against modified standards for state testing would be included in NAEP, with appropriate accommodations where necessary (NAGB, 2009).

The general messiness of the test data for this subgroup indicates a need to tighten and systematize requirements for reporting test data for students with disabilities. Rules for reporting test results for students with disabilities under NCLB might be clarified by requiring states to report one set of results for regular state tests and another set of results for all alternate assessments. Other approaches could be tried, such as allowing states to exempt students with significant cognitive disabilities (under the 1% rule) from state testing or to continue testing these students for reporting purposes but exempt them from AYP calculations. Results from alternate assessments aligned to modified standards (under the 2% rule) could still be included in regular state data reporting and used for AYP purposes, as long as it could be shown that both tests are indeed aligned to the same grade-level standards. In any case, the federal and state governments should take steps to clarify how results for students with disabilities on state tests should be reported so that trends will be more valid and meaningful in the future.

References

- Advocacy Institute. (2007). *Seven million students with disabilities need No Child Left Behind*. Retrieved on October 1, 2009, from www.advocacyinstitute.org/NCLB/AIWhySWDsNeedNCLB.pdf
- Center on Education Policy. (2004). *From the capital to the classroom: Year 2 of No Child Left Behind*. Washington, DC: Author.
- Center on Education Policy. (2005). *From the capital to the classroom: Year 3 of No Child Left Behind*. Washington, DC: Author.
- 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. (2009c). *State test score trends through 2007-08, part 3: Are achievement gaps closing and is achievement rising for all?* Washington, DC: Author.
- Consortium for Citizens with Disabilities. (2007, September 10). Testimony of Katy Beh Neas, CCD education task force co-chair, before the Committee on Education and Labor, U.S. House of Representatives, Washington, DC. Retrieved on October 1, 2009, from www.aucd.org/template/page.cfm?id=491
- Government Accountability Office. (2009). *No Child Left Behind: Enhancements in the Department of Education's review process could improve state academic assessments* [GAO 09-911]. Retrieved on October 24, 2009, from www.gao.gov/new.items/d09911.pdf
- Kitmitto, S., & de Mello, V.B. (2008, November). *Measuring the status and change of NAEP state inclusion rates for students with disabilities*. National Center for Education Statistics. Retrieved on November 5, 2009, from <http://nces.ed.gov/nationsreportcard/pdf/studies/2009453.pdf>
- Koretz, D.M., & Barton, K. (2003). *Assessing students with disabilities: Issues and evidence*. National Center for Research on Evaluation, Standards and Testing, University of California at Los Angeles. Retrieved on November 5, 2009, from www.cse.ucla.edu/products/Reports/TR587.pdf
- National Assessment Governing Board. (2009, July 22). *Technical advisory panel on uniform national rules for NAEP testing of students with disabilities: Report to the National Assessment Governing Board*. Retrieved on October 3, 2009, from <http://nagb.org/newsroom/PressReleasePDFs/SD-Panel-Report.pdf>
- National Center on Educational Outcomes. (2009, July). *Annual performance report: 2006-07 state assessment data*. Retrieved on October 3, 2009, from www.cehd.umn.edu/NCEO/OnlinePubs/APRreport2006-2007.pdf
- National Disability Rights Network. (2007). 2007 legislative agenda. Retrieved on October 1, 2009, from www.ndrn.org/policy/leg_LONG.pdf
- Quenemoen, R. (2006, July 12). Testimony before the Committee on Education and the Workforce, U.S. House of Representatives, Washington, DC. Retrieved on October 1, 2009, from www.cehd.umn.edu/nceo/OnlinePubs/EdWorkforceTestimony2006.htm
- Quenemoen, R. (2009). Students with disabilities: Expectations, academic achievement, and the critical role of inclusive standards-based assessments in improving outcomes. In L. Pinkus (Ed.), *Meaningful measurement: The role of assessments in improving high school education in the 21st century*. Retrieved on October 10, 2009, from www.all4ed.org/publication_material/reports/meaningfulmeasurement
- Stullich, S., Eisner, E., & McRary, J. (2007). *National assessment of title I: Final report, vol. I, implementation*. Washington, DC: U.S. Department of Education, Policy and Program Studies Service.

- Thompson, S. J., Johnstone, C. J., Thurlow, M. L., & Altman, J. R. (2005). *2005 state special education outcomes: Steps forward in a decade of change*. National Center on Educational Outcomes. Retrieved on October 3, 2009, from <http://education.umn.edu/NCEO/OnlinePubs/2005StateReport.htm/>
- Thurlow, M. (2004, March 3). NCLB and the inclusion of students with disabilities in accountability and assessment systems. Testimony before the Committee on Education and the Workforce, U.S. House of Representatives, Washington, DC. Retrieved on October 3, 2009, from www.harborhouselaw.com/articles/highstakes.nclb.inclusion.thurow.htm
- Thurlow, M., Altman, J., & Vang, M. (2009.) *Annual performance report: 2006-2007 state assessment data*. National Center on Educational Outcomes. Retrieved on October 21, 2009, from www.cehd.umn.edu/NCEO/OnlinePubs/APRreport2006-2007.pdf
- U.S. Department of Education. (2003, December 9). Title I—Improving the academic achievement of the disadvantaged; final rule [34 CFR 200]. *Federal Register*, pp. 68698-68708.
- U.S. Department of Education. (2007, April 9). Title I—Improving the academic achievement of the disadvantaged; final rule [34 CFR 200]. *Federal Register*, pp. 17747-17781.
- U.S. Department of Education, National Center for Education Statistics. (2009). *Digest of education statistics, 2008* (NCES 2009-020). Retrieved on October 29, 2009, from <http://nces.ed.gov/fastfacts/display.asp?id=64>

Credits and Acknowledgments

This report was written by Naomi Chudowsky and Victor Chudowsky, CEP consultants, who also conducted research for this study. Nancy Kober, CEP consultant, edited the report and assisted with the writing. 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 November 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

w: www.cep-dc.org

www.cep-dc.org