American Federation of Teachers, 2008

Sizing up state standards 2008*

Focus

Judges the clarity and specificity of states’ academic content standards in light of NCLB accountability requirements.

Methodology

Examined state content standards in four core content areas (English, math, science and social studies) posted on state Web sites as of October 2007 for all 50 states and D.C. Analyzed the standards using a set of criteria intended to determine whether standards contained enough information about what students should learn to provide the basis for coherent curricula and assessments.

Major Findings

- **Lack of strong standards in most states.** Only one state, Virginia, met the AFT criteria for strong standards in all levels and subjects. While some states have a lot of work ahead of them to improve their standards, others have to focus on only a few grades in one subject area.

- **Better standards in math and science.** Since 1995, states have done a better job with their standards in math and science than in English and social studies. Twenty-four states have strong math standards, and 22 have strong science standards. Only 8 states have strong English standards at all levels, and only 2 have strong social studies standards at all levels.

- **Variation by level.** For most subjects, middle school standards are the strongest, while high school standards are the weakest. The weaknesses at the high school level, in many cases, are due to the high school standards being clustered (e.g., one set of standards for grades 9-12) instead of being grade- or course-specific.

- **Reasons for weaker standards.** Standards failed to meet AFT criteria for three main reasons: They were repeated, clustered, or had missing or vague content. All three of these problems have the same consequences. Teachers do not have a common understanding of what students should have learned in the previous grade, what they are expected to master in the current grade, or what they are preparing them to learn in the following grade.

- **Early grade standards.** Too many states have clustered K-2 standards or have chosen not to write them at all. Nine states have clustered or no standards for K-2 in literacy and numeracy. Specific, coherent, grade-by-grade standards at the early grades are essential to building students’ background knowledge and vocabulary.

Where to Obtain


*For additional information on this topic, see the summary of the 2006 AFT report, Smart Testing: Let’s Get It Right.
American Federation of Teachers, 2006

Smart testing: Let’s get it right. How assessment-savvy have states become since NCLB?*  

Focus  

Examines whether states’ academic content standards are clear and specific for each subject and grade tested and whether state assessments in reading, math, and science are aligned with strong standards in light of NCLB’s emphasis on test-based accountability.

Methodology  

Reviewed 861 content standards documents (357 in reading, 357 in math, and 147 in science) posted on the state Web sites for all states and D.C in 2006. Also reviewed information about 833 state tests (357 in reading, 357 in math, and 119 in science).

Major Findings  

• **Alignment of standards and tests.** Seventy-four percent of the content standards across states met AFT’s criteria for strong standards, but only 52% of states’ tests were aligned to strong standards. States had done a better job of developing content standards than using them to drive assessment. Testing unaligned to strong standards was driving many accountability systems in 2006. As a result, the testing systems in many states were not yet “smart” enough to bear the weight of the accountability functions they were being asked to serve.

• **Specificity of standards.** Overall, content standards were more specific in 2006 than when AFT reviewed standards in 2001. This was particularly notable given that states had to develop 17 different sets of content standards after enactment of NCLB, versus the 9 required in 2001.

• **Weaknesses.** Some states did not post any information on their Web sites about which standards were being assessed by the tests they administered. States continued to struggle with articulating strong reading standards, including helpful descriptions of what students need to know at each grade level. States also had not articulated clearly enough what high school students should be learning in reading, math, and science. Nearly half of the testing (48%) was based on weak standards and/or unaligned tests.

Where to Obtain  


*For additional recent information on this topic, see the summary of the 2008 AFT report, *Sizing Up State Standards.*
Arizona State University, Education Policy Research Unit, 2006

The accuracy and effectiveness of adequate yearly progress, NCLB’s school evaluation system

Focus

Considers how well adequate yearly progress is working as the key element of the NCLB accountability system.

Methodology

Analyzed evidence from a variety of studies about major aspects of AYP.

Major Findings

- **Lack of evidence for AYP requirements.** In its 2006 form, AYP as an indicator of academic performance was not supported by reliable evidence. Nor was there any evidence at the time of this study that AYP demands, by themselves, will adequately provide schools serving children in poverty with the facilities, learning resources, qualified staff, or community support services needed to improve achievement.

- **100% proficient unrealistic.** Test score gains that may be attributable to the AYP process are modest and insufficient to achieve the goal of 100% proficiency by 2014.

- **Insufficient funding.** Whether conceived as implementation costs or remedial costs, NCLB has been significantly underfunded in a way that disproportionately penalizes schools attended by the neediest children.

- **Curriculum narrowing.** The coupling of high-stakes consequences with test scores increasingly appears to narrow curriculum in the opinion of state and local educators.

- **Flexibility measures not a solution.** Growth models and other “flexibility” measures do not have sufficient power to resolve the underlying problems of the AYP system. For example, all students must still meet the same standards by 2014 under a growth model. If the minimum size for disaggregated groups is reduced, the sum effect would be to make an already unworkable system less workable.

- **Recommendation to suspend AYP.** The report recommends that AYP sanctions be suspended until the premises underlying them can be either confirmed or refuted by solid, scientific research and unintended, negative consequences can be avoided.

Where to Obtain

Arizona State University, Education Policy Research Unit, 2005

High-stakes testing and student achievement: Problems for the No Child Left Behind Act

Focus

Investigates whether the pressure of high-stakes testing has influenced students’ academic performance

Methodology

To measure the impact of high-stakes testing pressure on achievement and to account for the differences in testing pressure among the states, the researchers created a Pressure Rating Index (PRI), which sought to capture the amount of pressure or “threat” associated with performance on a particular test. The index was developed by reviewing state legal requirements, interviewing state officials, and consulting media sources. The degree of pressure associated with specific state tests was determined by groups of graduate-level education students. The researchers analyzed correlations between the PRI and NAEP results from 1990 to 2003 in 25 states and used the PRI to replicate previous research.

Major Findings

- **No important influence.** Pressure created by high-stakes testing has had almost no important influence on student academic performance.

- **Greater test pressure in states with more minority students.** States with greater proportions of minority students have implemented accountability systems that exert greater pressure. This finding suggests that problems linked to high-stakes testing will disproportionately affect minority students.

- **Greater test pressure and dropout rates.** High-stakes testing pressure is negatively associated with the likelihood that 8th and 10th graders will still be in school in 12th grade. Increased testing pressure appears to be related to larger numbers of students being held back or dropping out of school.

- **No link between pressure and NAEP reading scores.** Increased testing pressure produced no gains in NAEP reading scores for students in grades 4 or 8.

- **Weak link in grade 4 math.** After 1996, increased testing pressure was weakly linked to later gains in NAEP math achievement at grade 4 for all ethnic subgroups. In grade 8 math, however, the study found no evidence that earlier pressure increases produced later achievement gains. The study authors noted that since math in the primary grades is far more standardized across the country than the math curriculum in middle school, drilling students and teaching to the test, as well as increased test pressure, could have played a role in these achievement gains.

Where to Obtain

http://epsl.asu.edu/epru/documents/EPSL-0509-105-EPRU.pdf
The Aspen Institute, Commission on No Child Left Behind, 2007

Children with disabilities and LEP students: Their impact on the AYP determinations of schools

**Focus**

Examines the impact of minimum subgroup size requirements in five states on determinations of adequate yearly progress for students with disabilities and English language learners.

**Methodology**

Collected student achievement data from school year 2004-05 for all schools in California, Florida, Michigan, Georgia, and Pennsylvania.

**Major Findings**

- **Impact of minimum subgroup size.** State-set requirements for minimum \( n \) size—the number of students that must be in a subgroup for that subgroup to count for AYP purposes—can have a considerable impact on the AYP status of a school. Since 2002, the general trend has been for states to increase their \( n \) sizes. Consequently, fewer schools are being held accountable for the performance of specific subgroups because these subgroups fall below their states’ minimum \( n \) size.

- **Schools with SWD and ELL subgroups too small to count.** The percentages of schools in the five states that did not have to report AYP determinations for the subgroup of students with disabilities ranged from 41% in Michigan to 92% in California. The percentages that did not have to report AYP determinations for the English language learner subgroup ranged from to 56% in California to 99% in Pennsylvania.

- **Students with disabilities and ELLs not sole reason for missing AYP.** Findings from the five states studied raise questions about claims that many schools do not make AYP solely based on performance of students with disabilities and ELLs. Even when these subgroups fell short of AYP targets, they were very often not the sole reason why a school failed to make AYP.

- **State data.** Specific data from the five states are shown in the two tables below:

### Impact of Students with Disabilities (SWDs) on Adequate Yearly Progress

<table>
<thead>
<tr>
<th>State &amp; ( n ) size*</th>
<th>% of all schools reporting AYP for subgroup</th>
<th>% of all schools that missed AYP in subgroup</th>
<th>% of schools not making AYP that missed solely due to subgroup</th>
<th>% of all tested SWDs enrolled in schools reporting AYP for subgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA – 100, or 50 and 15%</td>
<td>9%</td>
<td>4%</td>
<td>1%</td>
<td>28%</td>
</tr>
<tr>
<td>FL – 30 and 15%</td>
<td>58%</td>
<td>22%</td>
<td>2%</td>
<td>83%</td>
</tr>
<tr>
<td>GA – 40 or 10%</td>
<td>53%</td>
<td>10%</td>
<td>38%</td>
<td>80%</td>
</tr>
<tr>
<td>MI – 30</td>
<td>60%</td>
<td>3%</td>
<td>12%</td>
<td>70%</td>
</tr>
<tr>
<td>PA – 40†</td>
<td>11%</td>
<td>6%</td>
<td>19%</td>
<td>41%</td>
</tr>
</tbody>
</table>

*Percentages in \( n \) size column refer to percentage of total school enrollment.
†State has special provisions for calculating AYP in schools with \( n \) size below 40.
## Impact of English Language Learners on Adequate Yearly Progress

<table>
<thead>
<tr>
<th>State &amp; (n) size*</th>
<th>% of all schools reporting AYP for subgroup</th>
<th>% of all schools that missed AYP in subgroup</th>
<th>% of schools not making AYP that missed solely due to subgroup</th>
<th>% of all tested ELLs enrolled in schools reporting AYP for subgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA – 100, or 50 and 15%</td>
<td>44%</td>
<td>22%</td>
<td>12%</td>
<td>87%</td>
</tr>
<tr>
<td>FL – 30 and 15%</td>
<td>23%</td>
<td>14%</td>
<td>1%</td>
<td>80%</td>
</tr>
<tr>
<td>GA – 40 or 10%</td>
<td>10%</td>
<td>1%</td>
<td>2%</td>
<td>67%</td>
</tr>
<tr>
<td>MI – 30</td>
<td>9%</td>
<td>&lt;1%</td>
<td>2%</td>
<td>45%</td>
</tr>
<tr>
<td>PA – 40†</td>
<td>1%</td>
<td>&lt;1%</td>
<td>0%</td>
<td>20%</td>
</tr>
</tbody>
</table>

*Percentages in \(n\) size column refer to percentage of total school enrollment.
†State has special provisions for calculating AYP in schools with \(n\) size below 40.

### Where to Obtain

Many states have taken a “backloaded” approach to NCLB’s goal of all students scoring “proficient”

Focus

Examines the annual measurable objectives for student achievement set by all 50 states and the District of Columbia. These objectives specify interim performance targets on the way toward reaching the NCLB goal of 100% of students performing at the proficient level on state tests by the end of school year 2013-14.

Methodology

Analyzed state accountability plans posted on the U. S. Department of Education Web site or information retrieved directly from state department of education Web sites. Data were retrieved in January 2008.

Major Findings

- **Backloaded approach.** Twenty-three states have “backloaded” their trajectories for reaching 100% proficiency. In other words, they expect smaller achievement gains in the earlier years of NCLB and much steeper gains in later years, as 2014 grows closer. Some states expect leaps of 10 or more percentage points per year in the later years.

- **Incremental approach.** Another 25 states and D.C. have adopted a more incremental approach that assumes steadier progress toward the 100% proficient goal.

- **Blended approach.** The two remaining states used blended trajectories that do not fit readily into the backloaded or incremental categories.

- **Implications of backloaded approach.** Schools and districts in backloaded states are likely to have more difficulty making AYP as time goes on and could see a rise in the number of schools identified for NCLB improvement.

Where to Obtain

[www.cep-dc.org](http://www.cep-dc.org)
Center on Education Policy, 2007

No Child Left Behind at five: A review of changes to state accountability plans

Focus

Examines changes to NCLB accountability plans that were approved by the U.S. Department of Education in 2006.

Methodology

Reviewed decision letters to states from the U.S. Department of Education posted on ED’s Web site between January 1 and December 31, 2006. The decision letters reported changes to state accountability plans that were approved by ED (but not changes that were denied).

Major Findings

- **Compliance with testing systems.** Many changes requested were related to meeting the 2005-06 deadline for implementing tests that fulfilled NCLB requirements. Many states had to change cut scores and AYP targets because they had introduced new testing programs or tests at new grade levels.

- **Deadline for AYP notification.** Many states did not meet the NCLB requirement to report the AYP status of schools and districts before the start of the school year (in this case, school year 2006-07). These delays created confusion about which schools would have to undertake the improvement steps and interventions required by NCLB.

- **Flexibility in AYP determinations.** ED approved changes in 2006 that in effect made it easier for schools and districts to demonstrate AYP. These included, among others, the adoption of confidence intervals, indexing systems, and more lenient policies for counting scores from retests.

- **Learning from other states.** The flexibility permitted by ED in 2006 does not break new ground. Rather, more states are copying changes that ED had already allowed in other states or are applying the adjustments and flexibility described in policy guidance issued by Secretary of Education Margaret Spellings in 2005.

Where to Obtain

http://www.cep-dc.org
Center on Education Policy, 2006

From the capital to the classroom: Year 4 of the No Child Left Behind Act

Focus

Comprehensive study that describes the federal, state, and local implementation and impact of various provisions of NCLB during school years 2004-05 and 2005-06. Summarized below are the study’s findings about accountability and AYP issues.

Methodology

Collected data through a survey of all 50 states, a nationally representative survey of 299 school districts, case studies of 38 geographically diverse districts and 42 schools, three national forums, and six special analyses of critical issues in implementing NCLB.

Major Findings

- **AYP and schools identified for improvement.** About 16% of all schools and 24% of all school districts did not make adequate yearly progress based on 2004-05 testing. For the 2005-06 school year, about 14% of Title I schools, or 6,748 schools, were identified for improvement, corrective action, or restructuring. This was similar to the percentage identified the previous year, although the schools were not always the same schools. About 13% of school districts were in improvement.

- **Impact of policy changes.** The stable numbers of schools in improvement were partly due to changes in federal and state rules that made it easier for districts and schools to make AYP. Examples include the use a statistical technique called confidence intervals to create a sort of margin of error for AYP calculations; the use of index systems to give credit for gains by lower-achieving students; and increases in the minimum number of students that must be in a subgroup for that subgroup’s test scores to count for AYP. Other federally-approved state policy changes may have resulted in more students being counted as proficient, such as testing some students with disabilities against modified or alternate standards and counting passing scores from students who retake a test they previously failed.

- **Greater impact on urban districts** Although all school districts have been affected by NCLB, urban districts have experienced the greatest effects. The majority (54%) of Title I schools identified for improvement for school year 2005-06 were located in urban districts, even though just 27% of Title I schools were located in urban districts. About 90% of the schools in restructuring, the most serious stage of NCLB’s sanctions, were in urban districts. The diversity and poverty of urban districts were major reasons for this disproportionate impact of NCLB. Some urban districts must make AYP for up to 10 racial-ethnic and demographic subgroups, while some rural districts have to show progress for just two subgroups (white and low-income students). In addition, urban districts, due to their size, must make AYP for dozens of schools, while a small district may have just one elementary, one middle, and one high school. Increases in states’ minimum subgroup sizes also help smaller districts more than larger ones.

Where to Obtain

http://www.cep-dc.org
Center on Education Policy, 2005

*From the capital to the classroom: Year 3 of the No Child Left Behind Act*

**Focus**

Comprehensive study that describes the federal, state, and local implementation and impact of various provisions of NCLB during school years 2003-04 and 2004-05. Summarized below are the study’s findings about accountability and AYP issues.

**Methodology**

Collected data through a survey of 49 states, a nationally representative survey of 314 school districts, case studies of 36 geographically diverse districts and 37 schools, three national forums, and four special analyses of critical issues in implementing NCLB.

**Major Findings**

- **Students with disabilities and ELLs.** States and districts most often cited the accountability requirements for students with disabilities and English language learners as their greatest challenge. Although ED relaxed the requirements somewhat for both subgroups during 2004, many state and district officials indicated that the changes did not go far enough and viewed the testing policies for these groups as unfair, unrealistic, inappropriate, or instructionally meaningless.

- **Unrealistic 100% goal.** Many state and local respondents saw the goal of 100% of students performing at proficient levels by 2014 as unrealistic and expressed doubt about their ability to meet state AYP targets.

- **Flexibility and challenges.** Changes in federal and state policies offered more flexibility in NCLB implementation, which made it easier for both schools and districts to make AYP, at least in the short term. But states and districts still reported problems with NCLB accountability requirements, such as conflicts with pre-existing state accountability systems and difficulties meeting the deadline for reporting schools’ AYP status before the beginning of the school year.

- **Large and urban districts.** Over time, schools identified for NCLB improvement have become more concentrated in very large school districts and in urban areas. These types of districts tend to be more diverse, with more subgroups of students that must demonstrate AYP.

**Where to Obtain**

[http://www.cep-dc.org](http://www.cep-dc.org)
States test limits of federal AYP flexibility

Focus

Examines changes to NCLB accountability plans approved by the U.S. Department of Education in 2005.

Methodology

Reviewed decision letters to states from the U.S. Department of Education posted on ED’s Web site between January and October, 2005. The decision letters reported changes to state accountability plans that were approved by ED (but not changes that were denied).

Major Findings

- **AYP flexibility.** In 2005, the U.S. Department of Education granted states additional flexibility in how they determined adequate yearly progress for NCLB. Many of the changes to state accountability plans approved in 2005 had the effect of making it easier for schools to make AYP, including the following:
  - Confidence intervals, a statistical technique intended to make allowances for natural fluctuations in test scores unrelated to student learning
  - Performance indexes that allow schools to get “partial credit” for improved performance of students below the proficient level
  - Retesting, which allows students to retake a different version of the same test and permits schools to use a student’s best score to count toward AYP
  - Increased minimum subgroup sizes, which mean that in many schools, subgroups do not get counted for AYP purposes

- **Additional flexibility.** In 2005, ED also permitted flexibility in other areas, such as allowing a few districts to provide supplemental educational services to students before offering them school choice and issuing new rules that allow more students with disabilities to be tested against modified or alternate standards. Other state changes in accountability plans made necessary adjustments in response to states’ difficulties in administering the law.

- **Less transparency and comparability.** Over time, changes in state accountability plans for NCLB have increased the complexity and variety of state AYP formulas, making it harder for parents and others to understand what it means if a school does not make AYP. In addition, constant changes in AYP formulas make it difficult to tell whether student achievement is really improving, because percentages proficient or percentages of schools missing AYP may not be truly comparable from year to year.

- **ED delays and limited information.** ED often delayed in posting decision letters to states on its Web site and provided very little information about requested changes that were rejected or the rationales for accepting or rejecting changes.

Where to Obtain

http://www.cep-dc.org
Focus

Reviews federal and state implementation of the NCLB accountability requirements and their impact on minority and low-income students.

Methodology

Examined trends in the number and types of schools identified for improvement under NCLB over five years (2002-03 to 2006-07). Compared the demographic characteristics of students in schools identified for improvement and those in schools that made adequate yearly progress in six states with large proportions of minority and low-income students, (Arizona, California, Georgia, Illinois, New York, and Virginia). Data were drawn from the school achievement report cards required by NCLB and other reports issued by state departments of education.

Major Findings

• **Shortcomings of AYP mechanism.** As it stood in 2006, the AYP mechanism used to identify poorly performing schools was not working effectively to encourage improvement in these schools and proficiency in all schools. AYP did not seem to serve as a preventative measure.

• **Disproportionate impact on minority and low-income students.** Schools most likely to be identified as needing improvement are highly segregated and enroll a disproportionate share of a state’s minority and low-income students. The NCLB sanctions are concentrated on schools serving disadvantaged and minority students, whether or not those schools are making progress that isn’t measured by AYP.

• **State accountability changes.** The changes that the U. S. Department of Education has approved to state accountability plans, which were intended to reduce the number of schools identified for improvement, have complicated understanding of what accountability means and have made comparisons from one year to the next meaningless.

Where to Obtain

http://www.civilrightsproject.ucla.edu/research/esea/NCLB_Policy_Brief_Final.pdf
The Civil Rights Project at Harvard University, 2006

The unraveling of No Child Left Behind: How negotiated changes transform the law

Focus

Documents changes states made to their accountability plans. Examines how these policy shifts have affected the meaning of accountability and who benefits (and loses) from them.

Methodology

Reviewed decision letters from the U.S. Department of Education to all 50 states outlining the changes approved by ED through December 2005.

Major Findings

- **Compromises in accountability plans.** In 2004 and 2005, ED made extensive compromises by allowing states to make a wide variety of changes in their accountability plans. These changes reflected a political strategy to respond to growing opposition to NCLB by providing relief from some of its provisions. Many of the changes simply reduced the number of schools and districts identified for improvement without requiring any educational improvement.

- **Types of changes.** Several of these policy changes affected how students with disabilities and English language learners were counted for accountability purposes, revised how states calculated test participation rates, and relaxed the highly qualified teacher requirements. Other state-initiated changes to accountability plans were negotiated on a state-by-state basis, with no guidelines about the types of changes states could request, no information on how the requests would be judged, and no guarantees that changes approved in one state would be approved in another. The state-initiated amendments were extensive, and many changed how states determined AYP. These changes have complicated the meaning of AYP and obscured the ability of states, districts, and schools to show improvements in student performance.

- **Lack of uniformity.** Since the number and kinds of changes that states have adopted are not uniform across states, accountability no longer has a common meaning across states or even within states.

- **Winners and losers.** There were clear winners and losers from the changes. Some of the changes, such as the change in the method used to identify districts for improvement, made it harder for some districts, primarily those serving minorities, to make AYP. Others, such as the changes in the highly qualified teacher requirements, benefited some regions of the country more than others.

Where to Obtain

http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/29/dd/3a.pdf
Focus

Evaluates the efficacy of NCLB’s test-based accountability system and discusses options for the law.

Methodology

Solicited chapters from noted education scholars about major aspects of NCLB’s performance-based accountability system.

Major Findings

- **Shortcomings of current system.** The NCLB accountability system is not based on hard evidence and does not provide information necessary to know how well students are actually performing or what to do to improve learning and instruction.

- **Achievement results not promising.** A comparison of findings from NAEP with state assessment results shows that federal accountability has not systematically improved reading and math achievement or reduced achievement gaps. The nation has not focused on the kinds of serious, long-term reforms that can actually produce gains and narrow gaps for minority students.

- **Limited state capacity.** States have limited capacity to meet NCLB requirements and intervene in low-performing schools on the scale demanded by the law.

- **Negative effect on reform.** Many NCLB provisions—including the highly qualified teacher definition and the testing and accountability regulations—impede school reform and make it more difficult for high schools serving low-income students to do their work.

- **Inconsistency between dropout rates and AYP.** Only about 40% of the nation’s high schools that have high dropout rates have been identified for NCLB improvement.

Where to Obtain

Council of Chief State School Officers, 2008

Statewide educational accountability systems under the NCLB Act—A report on 2008 amendments to state plans

Focus

Describes the types of amendments to state accountability plans that states sought to have approved by the U. S. Department of Education and the decisions by ED to approve or deny them. These amendments would change elements of the accountability systems that all states must have in place to comply with NCLB.

Methodology

Analyzed information provided voluntarily by states to the Council of Chief State School Officers about amendments requested to state plans during 2007-08. State information included copies of state documents, e-mails, and reports of telephone conversations. Analyzed subsequent decision letters from ED to the states posted on the ED Web site.

Major Findings

• **High number of requests.** In 2007-08, the number of states submitting amendments (49 states plus D.C.) and the number of substantive amendments proposed (about 275) was the highest in any single year since NCLB was enacted in 2002. Vermont was the only state that did not submit amendments. Almost all decision letters were posted on ED’s Web site by the end of July 2008.

• **Types of amendments.** For 2007-08, the most frequent amendment requests related to the following issues (number of states requesting are in parentheses):
  - Continuing or initiating flexibility to use “modified” student academic achievement standards in AYP decisions for certain students with disabilities (30)
  - Implementing science assessments (17)
  - Including formerly served students with disabilities in AYP calculations for the disabilities subgroup (15)
  - Defining and calculating graduation rates for AYP decisions (16)
  - Changing the minimum subgroup size to be counted for AYP purposes (11)
  - Modifying how new schools and reorganized schools are defined for AYP (10)
  - Modifying how AYP determinations are made for schools not covered by state assessments and for small schools (10)
  - Modifying or clarifying the methods for calculating AYP in general (10)
  - Permitting states to use a growth model to determine AYP (10)
  - Modifying starting points, annual measurable objectives, and interim goals on the way to reaching NCLB’s ultimate goal of 100% proficient
  - Modifying or clarifying how the achievement of English language learners is included in AYP determinations or measures of English language acquisition (10)

• **Denied requests.** A few states sought amendments that ED has consistently denied, such as limiting schools identified for improvement to those that missed AYP in the same subject and subgroup for two consecutive years, or targeting sanctions in identified schools only to those subgroups that missed AYP.

Where to Obtain

Focus

Discusses why the narrow test-based accountability systems developed for NCLB cannot determine whether schools are performing satisfactorily or support interventions that ensure improvement. Proposes an alternative system for holding schools accountable.

Methodology

Commissioned surveys to determine what value the public and elected representatives place on the various goals of a well-designed accountability system. Interviewed teachers around the country about the effects of accountability policies on their instructional practices. Reviewed accountability plans in public and private fields.

Major Findings

• **Goals of accountability system.** According to surveys of the public and elected representatives, more than half of the weight of an accountability system should be devoted to the academic goals of basic knowledge and skills, critical thinking, appreciation of the arts and literature, and acquisition of occupation-specific technical skills, while the balance should be devoted to citizenship, social skills, and other physical and emotional health behaviors. Holding schools accountable for math and reading test scores has created incentives for educators to pay less attention to other curricular areas and goals for which they are not held accountable.

• **Flaws of current accountability systems.** Accountability policies for NCLB and state systems ignore the normal variation in student abilities and include such fanciful proficiency definitions that even the highest-scoring countries don’t come close to realizing them. A test-based accountability system based on a fixed proficiency point leads to excessive concentration on students whose performance is slightly below that point and ignores those who are either above or far below it. It also creates incentives for educators to game the system.

• **Accountability in other sectors.** A review of accountability plans in sectors such as health care, job training, welfare, and criminal justice reveals that an accountability plan that relies primarily on quantitative short-term measures (similar to test scores) without substantial qualitative evaluation will corrupt any institution.

• **Supplementing NAEP.** The model of the National Assessment of Educational Progress in its early years suggests some elements of a sophisticated accountability system, in that it focused on long-term outcomes by assessing young adults as well as schoolchildren and it measured behavioral as well as cognitive results of schooling. But because NAEP’s complex sampling methodology can only generate results at the state level (or for very large urban districts), it cannot tell policymakers how individual schools or other youth institutions are contributing to student results. Supplementing this information requires actual inspection of schools and other institutions of youth development.

• **Models for school inspection.** American school boards have lost sight of their obligation to hold schools accountable for outcomes defined by the public through democratic procedures. Elements of a democratic school accountability system could be adapted from existing arrangements, such as the current system of school
accreditation, which is based on school inspections but does not presently provide an adequate means of accountability. Some other nations more successfully use school visitations and inspectorate models for accountability purposes.

- **Alternative accountability system.** To hold schools, districts, and other youth development institutions accountable, information from tests of basic skills should be combined with a wide array of information from other sources, including tests of reasoning and critical thinking, and professional visitations by experienced and qualified educators. The federal role in such a system should include a funding distribution mechanism to ensure that states with limited fiscal capacity and relatively large numbers of disadvantaged children have sufficient resources to generate the expected youth outcomes. It should also include a vastly expanded NAEP that could give state leaders the information they need to determine if their youth demonstrate balanced achievement in eight broad goal areas.

**Where to Obtain**

http://www.epi.org/publications/entry/books_grading_education
Focus

Reports numbers of schools nationally that failed to make adequate yearly progress according to NCLB criteria in school year 2007-08.

Methodology

Analyzed data from 47 states and the District of Columbia on the number of schools making AYP based on tests administered in school year 2007-08.

Major Findings

- **Number of schools failing to make AYP.** Almost 30,000 U.S. schools failed to make AYP in school year 2007-08, an increase of 28% over school year 2006-07 in states with comparable data.

- **Number of schools identified for improvement.** Half of the schools that failed to make AYP in 2007-08 missed their achievement goals for two or more years, putting almost one in five of the nation’s public schools in some stage of NCLB improvement. The number facing sanctions represents a 13% increase over school year 2006-07 for states with comparable data.

- **Restructuring schools.** Based on their 2007-08 performance, 3,559 schools—4% of all schools rated—are in the planning or implementation stages of restructuring for the current school year. This is double the number in that category a year ago.

- **State-by-state differences.** While the national data suggest a steady increase in the number of schools failing to make AYP, state-by-state results show that states’ policy decisions can skew the results. In South Carolina, for example, 80% of public schools failed to make AYP in 2007-08, the highest proportion of any state. The increase can be partly attributed to the addition of new schools being rated for AYP and to the state’s decision to set challenging standards.

- **Backloaded achievement targets.** Decisions by 23 states to set low achievement targets in the early years of NCLB and quickly-rising targets in later years have also contributed to sharp increases in the number of schools not making AYP. Those states assumed that they would be able to ramp up student achievement by 2007-08, but the AYP results don’t reflect that. California, for example, had a dramatic increase in the percentage of schools failing to make AYP, from 34% in 2006-07 to 48% in 2007-08.

Where to Obtain

http://www.edweek.org/ew/articles/2008/12/18/16ayp.h28.html?tkn=MOWFHlIDjD5lWO1CQ6vONfAXdB6SB7RhJ8u&print=1
**Education Sector, 2006**

*Hot air: How states inflate their educational progress under NCLB*

**Focus**

Discusses how states have used their standard-setting flexibility to inflate their educational progress under NCLB and to minimize the number of schools that face scrutiny under the law.

**Methodology**

Reviewed data submitted by state departments of education to the U.S. Department of Education. The data came from Consolidated State Performance Reports submitted in March 2006.

**Major Findings**

- **General finding.** A significant number of states have used flexibility allowed by ED to make it seem as if their schools are doing better than they really are. Some states have dramatically inflated their high school graduation rates, teacher qualifications, school safety, and many other aspects of school performance under NCLB.

- **AYP statistical manipulations.** States used “statistical games” to make their performance appear better than it is. These included setting large minimum subgroup sizes, using confidence intervals (essentially a “plus or minus band”) around proficiency targets and safe harbor calculations, and requiring school districts to fall short of AYP targets at all three grade levels in the same subject before they fail to make AYP.

- **Proficiency and testing differences.** State differences in academic content standards, types of tests, and cut scores on tests for proficiency have created large variations in percentages proficient among states.

- **Persistently dangerous schools.** In 2006, states asserted that only 28 of the nation’s 95,000 schools were “persistently dangerous,” a designation that gives students the right to transfer under NCLB. Only six states reported any persistently dangerous schools at all.

- **HOUSSE process.** Most states made it relatively easy for veteran teachers to demonstrate they met the NCLB “highly qualified” teacher requirements by checking off a series of activities under the state’s HOUSSE (high objective uniform state standard of evaluation) process.

**Where to Obtain**

[http://www.educationsector.org/usr_doc/Hot_Air_NCLB.pdf](http://www.educationsector.org/usr_doc/Hot_Air_NCLB.pdf)
Government Accountability Office, 2006

No Child Left Behind Act: Education assistance could help states better measure progress of students with limited English proficiency

Focus

Describes the extent to which English language learners are meeting academic progress goals, what states have done to ensure the validity of their academic assessments and English language proficiency assessments, and how the U.S. Department of Education is supporting states’ efforts to meet NCLB testing requirements for ELLs.

Methodology

Collected state data on the AYP status of ELLs for 2003-04 from 48 states and district-level data from 18 states with the greatest numbers of or growth in ELLs. Did in-depth studies of ELL assessment practices in five states (CA, NE, NY, NC, and TX) and interviewed officials from 28 states, testing companies, and ED. Reviewed documents, convened a group of assessment experts, and surveyed all the states by e-mail on native language assessments.

Major Findings

- **Progress of ELLs.** In nearly two-thirds of the 48 states with sufficient data, ELLs did not meet annual state proficiency targets for reading or math in school year 2003-04. Further, in most states, these students generally did not do as well as other subgroups on state elementary school math tests.

- **Academic tests.** Officials in the five states studied reported following generally accepted procedures to ensure their academic tests for ELLs were valid and reliable. However, the expert group questioned whether all states were assessing ELLs in a valid way, noting that some lacked technical expertise. According to ED peer reviews of assessments in 38 states, 25 states did not provide adequate evidence of validity or reliability. Most states offer test accommodations to ELLs, but research is lacking on which accommodations are effective in mitigating language barriers. Several states used native language or alternate assessments, which are costly to develop and not appropriate for all students.

- **Language proficiency tests.** Many states implemented new English language proficiency assessments in 2006 to meet NCLB requirements, so complete information on validity and reliability is not yet available. A 2005 ED-funded review of 17 English language proficiency tests found insufficient documentation of validity.

- **ED assistance.** ED has offered technical assistance to help states assess ELLs but has issued little written guidance about developing English language proficiency tests. Officials in one-third of the 33 states contacted wanted more guidance on this issue. States have some flexibility in how they assess ELLs, but state officials wanted additional flexibility to appropriately track the academic progress of these students.

Where to Obtain


*GAO also delivered testimony on this issue on March 23, 2007 (http://www.gao.gov/new.items/d07646t.pdf). The testimony noted that since the original report was published, ED has initiated a partnership with states and other organizations to support development of valid assessment options for ELLs.*
Government Accountability Office, 2006

No Child Left Behind Act: States face challenges measuring academic growth that Education’s initiatives may help address*

Focus

Discusses the number of states using growth models—a term that encompasses a variety of methods for tracking changes in test scores or proficiency levels over time. Also describes why growth models are used, how they can measure progress toward achieving key NCLB goals, and what challenges states face in using growth models to meet the law’s key goals.

Methodology

In 2006, GAO surveyed all states and conducted site visits in California, Massachusetts, North Carolina, and Tennessee. In Massachusetts and Tennessee, GAO analyzed student-level data from selected schools.

Major Findings

- **States with growth models.** As of March 2006, 26 states were using growth models as part of their state education accountability systems, and 22 were either considering or in the process of implementing them. Most states using growth models measured progress for schools and for student subgroups, and seven states also measured growth for individual students. States used growth models to target resources for students that needed extra help or to award teachers bonuses based on school performance.

- **Purpose of growth models.** Certain growth models are capable of tracking progress toward 100% proficiency by 2014 and progress in closing achievement gaps. Massachusetts used its model to set targets based on the growth that it expected from schools and subgroups. Tennessee’s model projects students’ test scores and whether students will be proficient in the future.

- **Challenges.** States faced challenges measuring academic growth, such as creating data and assessment systems to support growth models and having sufficient personnel to analyze and communicate results. The use of growth models to determine AYP may also challenge states to make sure that students in low-performing schools receive needed assistance.

- **Pilot program.** U. S. Department of Education initiatives may help states address these challenges. In November 2006, ED started a pilot project to allow states to use growth models to determine adequate yearly progress; these models had to meet specific criteria, such as being able to track progress of individual students. ED also provided grants to states to track individual test scores over time.

Where to Obtain


Government Accountability Office, 2005

No Child Left Behind Act: Education could do more to help states better define graduation rates and improve knowledge about intervention strategies

Focus

Examines the definitions of graduation rates that states use for NCLB purposes, the actions the U.S. Department of Education has taken to help states meet legal requirements, the factors that affect the accuracy of states’ graduation rates, and ED’s role in ensuring accurate data. Also discusses interventions that could increase graduation rates and assesses ED’s efforts to disseminate intervention research.

Methodology

Surveyed 49 states and Puerto Rico. Also conducted a case study in one state to calculate graduation rates, site visits in three states to review data accuracy, site visits in six states to observe interventions and interview staff, and telephone interviews in 20 states to get information about definitions used, implementation status, and ED guidance.

Major Findings

- **Cohort definition.** In 2005, 12 states determined graduation rates for purposes of NCLB by using the cohort definition, which tracks students from the time they enter high school until the time they leave. Thirty-two states used a definition based primarily on the number of graduates and the number of dropouts over a four-year period. The remaining states used other definitions. Because the cohort definition is more precise, most states that were not using it planned to do so once they developed the capability to track students over time, a capability that many state data systems did not have at the time of this study.

- **ED guidance.** ED assisted states with definitions of graduation rates primarily on a case-by-case basis, but had not provided guidance to all states on ways to account for selected students, such as students with disabilities. As a result, state approaches for calculating graduation rates were not consistent.

- **Accuracy.** Student mobility was the primary factor affecting the accuracy of graduation rates. Another factor was whether states verified student data; fewer than half the states conducted audits of data used to calculate graduation rates. ED had taken steps to help states address data accuracy issues, but ED officials said they could assess only those state systems that had been in place for a while.

- **Promising interventions.** Many interventions were being used to raise graduation rates, but few had been rigorously evaluated. Based on its state and school visits, GAO identified five interventions that had been rigorously evaluated and showed potential for improving graduation rates, such as Project GRAD. Other interventions in schools visited by GAO were considered promising by experts and officials.

- **Lack of ED dissemination.** As of 2005, the U.S. Department of Education had not acted on GAO’s 2002 recommendation that it evaluate intervention research, which ED agreed with, and had done little to disseminate this type of research.

Where to Obtain

Great Lakes Center for Education Research and Practice, 2005

The impact of the adequate yearly progress requirement of the federal “No Child Left Behind” Act on schools in the Great Lakes region

Focus

Examines the implementation of the NCLB accountability and adequate yearly progress requirements in the six states of the Great Lakes region (Illinois, Indiana, Michigan, Minnesota, Ohio, and Wisconsin).

Methodology

Analyzed the AYP status of schools in the Great Lakes states and the annual objectives schools must meet to make AYP.

Major Findings

- **Assumptions for steep increases in later years.** The annual objectives for four of the six Great Lakes states assumed incremental increases in the percentage of students performing at the proficient level for the early years of NCLB implementation. However, these are followed by steep, backloaded objectives that assume very large increases in performance in the later years.

- **Likelihood of widespread failure to make AYP.** None of the Great Lakes states was among the top ten states with the highest percentage of schools failing to make AYP in either 2003 or 2004. By 2014, however, approximately 85% of schools in the Great Lakes states were projected to fail to make AYP under the most optimistic scenarios. Under more realistic circumstances, the overall failure rate is projected to be at or above 95%.

- **Concerns about NCLB accountability.** The authors of the study questioned the sustainability of the AYP requirements and cautioned that schools are not capable of closing the achievement gap without resolving underlying social problems. Adequate funding for remediation and social infrastructure is essential to meeting the goals of NCLB.

Where to Obtain

Citation: RAND, 2007

Standards-based accountability under No Child Left Behind: Experiences of teachers and administrators in three states

Focus

Describes how NCLB-related accountability policies have influenced attitudes and actions at the district, school, and classroom levels in three states, with a focus on math and science.

Methodology

Analyzed data from large-scale surveys and interviews and small-scale case studies conducted during school years 2003–04, 2004–05, and 2005–06 in California, Georgia, and Pennsylvania. Information was gathered from teachers and superintendents in a representative sample of 27 school districts in the three states and in 125 randomly selected elementary and middle schools within those districts. Site visits were conducted in 14 schools in 2003-04 and 16 schools in 2004-05. The report focuses on 2004-05 data.

Major Findings

- **State differences.** State accountability systems developed in response to NCLB differed across the three states in terms of the content of their academic standards, difficulty of performance standards, methods for calculating adequate yearly progress, trajectories for reaching 100% proficiency, support and technical assistance mechanisms, and other aspects. Many of these differences were related to state policies in place before NCLB, such as the degree of state involvement in standards-based accountability.

- **School improvement activities.** Despite state differences, districts and schools in the three states responded to accountability actively and in broadly similar ways. The school improvement activities described as most important by superintendents included aligning curriculum with state standards and assessments, using data for decision making, and providing extra support to low-performing students. In addition to these actions, most superintendents reported providing technical assistance to help schools improve and offering professional development for principals and teachers; a large number of principals reported providing extra learning opportunities for low-performing students. Other common activities included implementing test preparation and giving interim tests. District actions appeared to be influenced by the specific content and features of state standards and assessments.

- **Teacher views.** Teachers noted that NCLB had influenced their instruction in some beneficial ways, including changes to align instruction with standards and improve their own practices. Teachers also described some less desirable responses to NCLB, such as narrowing curriculum and instruction toward tested topics or certain problem formats and focusing more on “bubble kids” who were already closest to reaching proficiency. Teachers also expressed concern about negative effects of accountability requirements on learning for high-achieving students.

- **Specific concerns.** Most superintendents, principals, and teachers expressed support for the idea of standards-based accountability, but these groups had different concerns about specific features of accountability systems. Most administrators, but only a small minority of teachers, thought state test scores accurately reflected student achievement. Administrators were more likely than teachers to think the pressure of accountability led to improved curriculum and
learning. Teachers expressed concerns about reduced morale and other negative effects on their teaching. Still, teachers reported an increased focus on achievement in their schools as a result of NCLB, as well as greater coordination of and rigor in the curriculum.

- **Adverse effects.** Both teachers and administrators identified factors they believed adversely affected their efforts to meet NCLB goals, including inadequate funding, insufficient numbers of highly qualified teachers, and insufficient instructional and planning time. Teachers said that students’ lack of basic skills, inadequate parent support, and student absenteeism and tardiness hampered their efforts to promote high achievement.

**Where to Obtain**

Thomas B. Fordham Institute, 2009

The accountability illusion

Focus

Examines the NCLB accountability systems and the basic adequate yearly progress rules as they operate in practice in 28 states.

Methodology

Evaluated student performance in 18 elementary schools in 28 states and 18 middle schools in 26 states; the schools varied by size, achievement, diversity, and other factors. Analyzed whether a school in a particular state would make AYP according to the proficiency cut scores and 2008 annual AYP targets of other states, and if not, what factors within NCLB explained the difference. As part of the AYP analysis, applied confidence intervals according to each state’s rules and evaluated the performance of all subgroups within a school that met or exceeded each state’s minimum subgroup size requirements.

Major Findings

- **Different results using different states’ AYP rules.** Within the elementary school sample, the number of schools that made AYP varied greatly by state. Almost all of the schools in the sample failed to make AYP in some states, and nearly all of these same schools made AYP in others. In Massachusetts, for example, a state with high proficiency cut scores and relatively challenging annual targets and AYP rules, only 1 of 18 elementary schools made AYP; in Wisconsin 17 schools made AYP.

- **More consistency but lower performance in middle schools.** There was more consistency in AYP results across states among the middle schools in the sample because so few of these schools made AYP in any state. In 21 of the 26 states studied, two or fewer middle schools made AYP. In no state did even half of the 18 middle schools meet the 2008 AYP requirements. This was mostly because middle schools tended to enroll more students than elementary schools and generally had subgroups that were large enough to count separately for accountability purposes. Although the racial-ethnic subgroups in the sample schools met annual targets in many states, the subgroups of students with disabilities and English language learners, when large enough to count, failed to meet AYP targets in almost all of the schools in nearly every state.

- **Importance of minimum subgroup size rules.** State rules for minimum subgroup size (called n size) were critical in determining whether subgroup performance would hurt a school’s chances of making AYP. For example, the highest performing middle school in the sample had strong performance overall and showed growth in students’ performance over time, but it failed to make AYP in 21 of the 26 states because of subgroup performance. The states where this school did make AYP had larger n sizes. Generally, the lower the state’s n size, the more subgroups for which the typical school is accountable, and the more separate targets a school must hit.

Where to Obtain

http://www.edexcellence.net/index.cfm/news_the-accountability-illusion
Thomas B. Fordham Institute, 2007

The proficiency illusion

Focus

Examines the relative rigor of states’ proficiency standards (cut scores students must reach on state tests to be considered proficient) and whether some states’ tests are harder to pass than others. Considers whether state proficiency standards have changed since NCLB was enacted and whether standards in different grades are equivalent in difficulty.

Methodology

Estimated the differences among states in proficiency cut scores by analyzing test data from schools in 26 states whose pupils participated in both state tests and assessments developed by the Northwest Evaluation Association (NWEA).

Major Findings

- **Wide variation among states.** State tests varied greatly in difficulty. By the study’s estimates, state cut scores for proficient performance ranged from the 6th percentile on the NWEA scale to the 77th percentile.

- **Changes in test difficulty.** Most state tests had not changed in difficulty in recent years. Still, cut score estimates in reading and/or math declined significantly in at least two grades in eight states and rose in just four states.

- **Differences between reading and math.** Math tests were consistently more difficult to pass than reading tests. In seven states, the difference in cut scores between 8th grade math and reading was more than 10 percentile points, which could create a misimpression that students are doing better in reading when they aren’t.

- **Cut scores and pass rates.** Improvements in student proficiency rates on state tests can largely be explained by declines in the estimated cut scores for those tests.

- **Differences in cut score by grade level.** Eighth-grade tests were consistently and dramatically more difficult to pass than those in earlier grades (even after taking into account obvious differences in subject-matter complexity and children’s academic development). In three states, differences between 3rd and 8th grade cut scores in reading were 20 or more percentile points. Across states, this situation could give a possibly false impression that elementary schools are performing at much higher levels than middle schools.

- **Lack of meaning of proficiency.** Five years into implementation of NCLB, there was no common understanding of what “proficiency” meant. Its definition varies from state to state, from year to year, from subject to subject, and from grade level to grade level. This suggests that the goal of achieving “100 percent proficiency” has no coherent meaning, either.

Where to Obtain

http://www.edexcellence.net/doc/The_Proficiency_Illusion.pdf
Thomas B. Fordham Institute, 2006

2006: The state of state standards

Focus

Rates the quality of state academic content standards in all states and examines the relationship of the standards ratings to state achievement on NAEP.

Methodology

Selected experts to review and give quality grades to the standards of the 50 states and the District of Columbia. Standards were examined in five subjects: U.S. history (2003), English/language arts (2005), mathematics (2005), science (2005), and world history (2005). Compared each state’s rating for its standards with its performance on NAEP.

Major Findings

- **Ratings of standards.** State standards as a whole were no better in 2006 than in 2000. The average grade given by the study’s standards raters was a C-minus, the same as in 2000. Two-thirds of U.S. students attended school in states with academic standards in the C, D, and F range. Still, several states have shown marked progress in the quality of their standards, especially Indiana, New York, Georgia, and New Mexico. But other states have made their standards worse, including Utah, Nebraska, New Hampshire, and Wisconsin.

- **Link between strong standards and NAEP gains.** According to several indicators, states with strong standards have made gains on the NAEP assessments.

- **Reading.** In grade 4 reading, 10 states made statistically significant gains from 1998 to 2005 in the percentage of students performing at the proficient level on NAEP. Nine of those 10 states received a grade of C or above from the raters for their English/language arts standards.

- **Science.** Five states made significant gains on the science NAEP between 2000 and 2005 at both the 4th and 8th grade levels. Three of the five states had the best science standards in the nation, according to the raters.

- **Math.** The relationship between the quality of standards and NAEP scores was less clear in mathematics, although four of the six states that received “honors” grades from the raters also posted statistically significant gains on the grade 8 NAEP from 2000 to 2005, either for the state as a whole or for their poor and minority students. However, many other states made progress, too.

Where to Obtain

http://www.edexcellence.net/detail/news.cfm?news_id=358&id=130

Mapping 2005 state proficiency standards onto the NAEP scales

Focus

Focuses on comparing the relative stringency of states’ academic standards by mapping state standards for proficient performance on state reading and mathematics tests onto the appropriate scoring scale of the National Assessment of Educational Progress (NAEP).

Methodology

Developed a complex methodology for mapping states’ cut scores for proficient performance (called “proficiency standards” in the report) at grades 4 and 8 for school year 2004-05 onto a common scale—the NAEP scoring scale. Compared the results of this mapping with states’ average scores on the 2005 NAEP. The analysis included 32 states for grade 4 reading, 34 states for grade 8 reading, 33 states for grade 4 math, and 36 states for grade 8 math.

Major Findings

- **Differences in stringency of state standards.** There is a strong negative correlation between the proportions of students meeting states’ proficiency standards and the NAEP score equivalents to those standards. In other words, states with higher percentages of students scoring proficient on state tests tend to have lower NAEP score equivalents when their proficiency standards are mapped onto the NAEP scale. This suggests that differences among states in percentages proficient can be largely attributed to differences in the stringency of their standards.

- **Comparison of state proficiency standards and NAEP scores.** There is, at best, a weak relationship between the NAEP score equivalents for a state’s proficiency standard and its average scores on NAEP.

- **Difference in state and NAEP definitions of proficient.** NAEP has its own cut scores, or standards, that define various achievement levels on its assessments. When state proficiency standards were mapped onto the NAEP scoring scale, the resulting score equivalents for most states fell below the NAEP cut score for “proficient” performance, and in many states fell below the NAEP cut score for “basic” performance. In other words, state definitions of “proficient” performance are often less challenging than the NAEP standard for proficient performance and are sometimes less challenging than the NAEP “basic” standard.

Where to Obtain

Focus

Comprehensive study that describes the progress of states, districts, and schools through school year 2004–05 in implementing key provisions of Title I. Summarized below are the study’s findings about accountability and adequate yearly progress.

Methodology

Drew on data from a set of implementation studies by the U.S. Department of Education. Data for these studies came from surveys conducted in a nationally representative sample of school districts, other state and local surveys, state performance reports, and the National Assessment of Educational Progress.

Major Findings

- **Assessments.** During 2005-06, all states administered reading and mathematics assessments intended to meet NCLB requirements. As of September 1, 2007, 24 state assessment systems had been approved as meeting all NCLB requirements through the U. S. Department of Education’s peer review process. Most states had met the requirement to annually assess at least 95% of their students, although 15 states had not met the test participation requirement for one or more subgroups.

- **Identification for improvement.** For school year 2005-06, 12% of the nation’s schools (11,648 schools) were identified for improvement based 2004-05 testing. Title I schools accounted for 84% of identified schools. About 32% of high-poverty schools and 31% of high-minority-enrollment schools were identified for 2004-05, compared with 4% of schools with low concentrations of these students. Urban schools were also more likely to be identified than other schools.

- **Adequate yearly progress.** Three-fourths of all schools and districts met all applicable AYP targets in 2004-05 testing. Schools most commonly missed AYP targets for the achievement of all students and/or multiple subgroups. Of the schools that did not make AYP, only 21% of schools missed due to a single subgroup (most often students with disabilities). Schools in states with more challenging proficiency standards (relative to NAEP) were less likely to make AYP and had much further to go to reach the goal of 100% proficient by 2014 than schools in states with lower standards. Schools that were held accountable for more subgroups were less likely to make AYP. For example, 45% of schools that had to calculate AYP for six or more subgroups did not make AYP, compared with 5% of schools with just one subgroup.

- **Pace toward 100% proficient.** Slightly more than half the states have set “delayed acceleration” trajectories for the pace of improvement schools are expected to make to reach the 100% proficient goal. Rather than expecting schools to improve in roughly equal increments each year, these states expect much more rapid growth from 2009 to 2014 than from 2004 through 2009.

Where to Obtain


*For additional information about topics covered in this study, see the description for U. S. Department of Education, State and local implementation of the No Child Left Behind Act. Vol. III, interim report.*
Focus

Describes state and local implementation of the standards, assessment, and accountability provisions of Titles I and III of ESEA (as amended by NCLB) through school year 2004–05.

Methodology

Analyzed state data and documents; interviewed state officials; and conducted a survey of a nationally representative sample of 300 school districts and of 1,483 elementary, middle, and high schools in those districts. Individuals surveyed included more than 10,000 teachers, 1,783 principals and district administrators, and 950 Title I paraprofessionals.

Major Findings

- **State progress.** States, districts and schools had mostly met the relevant NCLB accountability requirements through 2004–05. All states, D.C., and Puerto Rico had adopted achievement standards in reading and math and other required indicators.

- **States behind schedule.** Twenty states were behind schedule in implementing assessments of English language proficiency. And 20 states were unable to notify schools, even preliminarily, about their test performance before September 2004.

- **State variations.** State accountability systems vary significantly in such areas as achievement levels required for proficiency, types of assessments, and pace of improvement expected to reach 100% proficient in 2014.

- **Disabilities and ELL subgroups.** About one-third of the schools that did not make AYP in 2003-04 missed targets for students with disabilities or English language learners. About two-thirds of the schools that missed targets for these subgroups reported needing technical assistance to improve instruction for these students.

- **Types of schools identified for improvement.** Schools identified for improvement were most likely to be high-poverty, high-minority-enrollment, large urban schools—the schools to which Title I has historically directed substantial resources.

- **Improvement efforts.** Nearly all schools reported making multiple improvement efforts. Schools identified for improvement focused on more areas of improvement than non-identified schools. Schools also reported receiving technical assistance that met their needs, except that about half of the schools needing assistance in serving students with disabilities or ELLs did not have these needs met. States and districts were implementing the required interventions in schools identified for improvement and corrective action but were not implementing the required actions in most of the 1,199 schools in restructuring.

Where to Obtain


*For additional and updated information about topics covered in this study, see the description for U. S. Department of Education, National Assessment of Title I: Final report, Vol. 1.*