Statewide Educational Accountability Systems Under the NCLB Act—A Report on 2008 Amendments to State Plans

A Summary of State Requests in 2007-08 for Amendments to Their Educational Accountability Plans Under the No Child Left Behind Act of 2001 (NCLB)

Willam J. Erpenbach

A report commissioned by the CCSSO Accountability Systems and Reporting State Collaborative - SCASS
COUNCIL OF CHIEF STATE SCHOOL OFFICERS
The Council of Chief State School Officers (CCSSO) is a nonpartisan, nationwide, nonprofit organization of public officials who head departments of elementary and secondary education in the states, the District of Columbia, the Department of Defense Education Activity, and five U.S. extra-state jurisdictions. CCSSO provides leadership, advocacy, and technical assistance on major educational issues. The Council seeks member consensus on major educational issues and expresses their views to civic and professional organizations, federal agencies, Congress, and the public.

STATE COLLABORATIVE ON ASSESSMENT AND STUDENT STANDARDS
The State Collaborative on Assessment and Student Standards (SCASS) Project was initiated in 1991 to encourage and assist States in working collaboratively on assessment design and development of high quality assessments in relation to standards for student learning. State education agencies voluntarily support collaborative work across a variety of topics and subject areas. The Accountability Systems and Reporting (ASR) state collaborative project has been assisting States to develop and improve statewide systems since 2000. Currently 28 state education agencies are participating members, together with researchers, consultants, and measurement specialists. Priority activities of the ASR Collaborative are: state planning and development to meet requirements for state systems under NCLB, sharing strategies and research-based evidence to improve the validity of accountability systems, developing reporting formats that improve information access for constituents, and establishing standards for the essential components of accountability systems.

Funding support for this paper was entirely from member States of the Accountability Systems and Reporting State Collaborative (ASR-SCASS). For further information about ASR SCASS and resources on accountability across the 50 States, see the CCSSO website: http://www.ccsso.org/projects/Accountabilty_Systems/.

2008
Council of Chief State School Officers
Rick Melmer (South Dakota), President
T. Kenneth James (Arkansas), President-Elect
Elizabeth Burmaster (Wisconsin), Past President

Gene Wilhoit, Executive Director

Rolf K. Blank, Director of Education Indicators

Council of Chief State School Officers
Attn: Publications
One Massachusetts Ave., NW, Suite 700
Washington, DC 20001
202-336-7016
Fax: 202-408-8072
www.ccsso.org

Copyright © 2008 by the Council of Chief State School Officers, Washington, DC.
All rights reserved.
Statewide Educational Accountability Systems Under the NCLB Act—A Report on 2008 Amendments to State Plans

A Summary of State Requests in 2007-08 for Amendments to Their Educational Accountability Plans Under the No Child Left Behind Act of 2001 (NCLB)

A report commissioned by the CCSSO Accountability Systems and Reporting State Collaborative - SCASS

William J. Erpenbach
October 2008
Acknowledgements

The author wishes to acknowledge the support and assistance of the ASR SCASS member states and individuals in every State Department of Education who so graciously contributed answers and insights reflected in this paper. The author also recognizes the contributions of Rolf Blank, CCSSO ASR SCASS coordinator, for his guidance of the paper and interaction with state accountability coordinators. Special recognition also goes to Nina de las Alas, CCSSO associate, for her expert assistance and perseverance in tracking down countless details to make sure that “we had it right” in chronicling the many, many accountability workbook amendments states sought and the resultant decisions by the U. S. Department of Education (ED). Finally, special thanks go to Abigail Potts at ED for helping us “track down” decision letters in the most expeditious manner.
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of Equi-percentile Adjustments for AYP Determinations</td>
<td>40</td>
</tr>
<tr>
<td>Exempting Schools and Districts Not Receiving Title I Funds</td>
<td>41</td>
</tr>
<tr>
<td>Adequate Yearly Progress (AYP) Indicators</td>
<td>41</td>
</tr>
<tr>
<td>Percent Proficient—Use of Indexing</td>
<td>41</td>
</tr>
<tr>
<td>Participation Rate</td>
<td>43</td>
</tr>
<tr>
<td>Other Academic Indicators</td>
<td>44</td>
</tr>
<tr>
<td>Growth Models</td>
<td>49</td>
</tr>
<tr>
<td>Annual Measurable Objectives and Intermediate Goals</td>
<td>51</td>
</tr>
<tr>
<td>Strategies to Enhance Reliability</td>
<td>52</td>
</tr>
<tr>
<td>Uniform Averaging</td>
<td>53</td>
</tr>
<tr>
<td>Minimum n</td>
<td>53</td>
</tr>
<tr>
<td>Confidence Intervals and Standard Errors of Measurement</td>
<td>55</td>
</tr>
<tr>
<td>Safe Harbor</td>
<td>55</td>
</tr>
<tr>
<td>Inclusion of All Students in Accountability</td>
<td>58</td>
</tr>
<tr>
<td>Inclusion of Students with Disabilities in Assessments and AYP</td>
<td>58</td>
</tr>
<tr>
<td>Modified Achievement Standards and the 2% Solution</td>
<td>58</td>
</tr>
<tr>
<td>Defining the Students with Disabilities Student Group for AYP Decisions</td>
<td>63</td>
</tr>
<tr>
<td>Changing AYP Measures for the SWDs Student Group</td>
<td>64</td>
</tr>
<tr>
<td>Inclusion of English Language Learners in Assessments and AYP</td>
<td>65</td>
</tr>
<tr>
<td>Inclusion Policies Generally</td>
<td>66</td>
</tr>
<tr>
<td>AYP Policies Related to the ELL Student Group</td>
<td>67</td>
</tr>
<tr>
<td>Use of Alternate Assessments for ELLs</td>
<td>71</td>
</tr>
<tr>
<td>AYP Consequences, Reporting, and Appeals</td>
<td>71</td>
</tr>
<tr>
<td>Public School Choice and Supplemental Educational Services</td>
<td>71</td>
</tr>
<tr>
<td>Changing the Order of SES and Choice Sanctions</td>
<td>72</td>
</tr>
<tr>
<td>Targeting SES and Choice</td>
<td>73</td>
</tr>
<tr>
<td>School Districts Identified for Improvement as SES Providers</td>
<td>74</td>
</tr>
<tr>
<td>Reporting</td>
<td>75</td>
</tr>
<tr>
<td>Appeals to Identification for Improvement Determinations</td>
<td>76</td>
</tr>
<tr>
<td>The Next Chapter</td>
<td>77</td>
</tr>
<tr>
<td>Chapter 3: Observations and Conclusions</td>
<td>78</td>
</tr>
<tr>
<td>Unanticipated Approvals</td>
<td>78</td>
</tr>
<tr>
<td>Consistently Denied Requests</td>
<td>79</td>
</tr>
<tr>
<td>Next Steps</td>
<td>79</td>
</tr>
<tr>
<td>References and Other Resource Material</td>
<td>81</td>
</tr>
<tr>
<td>APPENDIX A – Acronyms Used In This Paper</td>
<td>89</td>
</tr>
<tr>
<td>APPENDIX B – Timeline of Significant Events Related to State Accountability Plans Under NCLB</td>
<td>90</td>
</tr>
<tr>
<td>APPENDIX C – State Educational Accountability Workbook Decisions by the U. S. Department of Education that Changed Between 2003 and 2008</td>
<td>93</td>
</tr>
</tbody>
</table>
Executive Summary

Conclusion of the 2007-08 school year marks six full years of NCLB’s implementation into the nation’s public schools and school districts. Throughout this period, numerous states—sometimes almost every one—have annually sought to amend the educational accountability systems they developed in 2002 consistent with the law’s requirements and goals to improve teaching and student achievement. In most cases, the proposed amendments were initiated by the individual states. However, some accountability workbook amendments were submitted at the direction of the U. S. Department of Education (ED) as a result of a finding of non-compliance or other technicality usually resulting from a Title I Monitoring visit.

Once again this year, the number of states submitting amendments and the number of substantive amendments proposed was quite surprising, the highest in a single year since 2002 when accountability system workbooks were first required to be submitted for peer review—49 states and the District of Columbia and approximately 275 substantive changes.1 Vermont was the only state that did not submit educational accountability workbook amendments.

Last year, all but six states submitted formal amendment requests, ranging from one to ten or more per state. For the first time, every state requesting an amendment received a decision letter last year—as evidenced by postings (almost all letters were posted by the end of July) at ED’s website, http://www.ed.gov/admins/lead/account/letters/index.html. This was a significant change over past years. And, for the second consecutive year, ED again sent decision letters to every state and the District of Columbia albeit on a slower timeline. While most of the decision letters had been issued by mid-August the final letters were not released until late September.

For 2007-08, the most frequent amendment requests related to:

- 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 SWDs in subgroup AYP calculations (15);
- Defining and calculating graduation rates for AYP decisions (16);
- Changing minimum n’s (11);
- Modifying the manner in which new schools and reorganized schools are defined (10);
- Modifying the manner in which AYP determinations are made for schools not covered by state assessments and for small schools (10);
- Modifying or clarifying the manner in which AYP is calculated generally (10);
- Permitting states to use a growth model (10);
- Modifying starting points, AMOs, and IGs; and
- Modifying or clarifying the manner in which the achievement of LEP students is included in AYP determination or in English language acquisition measures (10).

Below, a summary is presented of the accountability workbook amendments states sought in 2008. A few states continued to seek amendments that ED has consistently denied in prior years, such as AYP

---

1 In late 2007, Elementary and Secondary Education Assistant Secretary Kerri Briggs sent chief state school officers a letter in which she stated, “Given that Congress is actively working through the reauthorization process, we would not expect to receive major amendments to States’ accountability plans at this time” (Briggs, 2007, November 19).
decisions based on missing the same target two years in a row. Some requests simply involve adoption of amendments previously approved for other states. Not included in this summary, or elsewhere, are the many minor requests that states submitted to update their workbooks.

The appendices include a list of acronyms used in this paper (Appendix A), a timeline of the significant events related to implementation of NCLB since its enactment in January 2002 (Appendix B), and a table listing changes in ED policy that have occurred since 2002 (Appendix C).

States’ requested amendments for 2007-08 AYP determinations included the following:

**Standards and Assessments**

**Changes to Assessment Systems**
- Revise schedule of or implement new statewide assessments: **Illinois, New Mexico, and West Virginia**
- Use “substitute assessments” for students enrolled in the “innovative schools:” **North Carolina**
- Switch from high school assessments to end-of-course assessments at this level: **Missouri**
- Implement new science assessments or otherwise add science assessments in the grade spans 3-5, 6-9, and 10-12 to educational accountability workbooks: **Alabama, Alaska, Arizona, District of Columbia, Hawaii, Illinois, Kansas, Kentucky, Maine, Maryland, Michigan, Missouri, North Carolina (2), Pennsylvania, Utah, Virginia, and Wisconsin**
- Revise an alternate assessment for significantly cognitively disabled students: **Colorado**
- Discontinue use of portfolio assessments or make other changes related to LEP students (see also inclusion): **Illinois**

**Use of ACT, SAT, AP, or IB Exams as High School Accountability Assessments**
- Substitute use of Advanced Placement/International Baccalaureate science assessments for the NCLB high school science assessment: **Maryland**
- Use SAT as high school assessment: **Maine**

**Banking of Test Results**
- Bank high school assessment results: **Wyoming**

**Recalculate AYP Using “Best Scores” or Results from Subsequent Test Administrations**
- Use “best score” for AYP calculations when students re-take state tests: **Georgia** (beginning in 2008-09), **Maryland**, and **Tennessee**

**Miscellaneous Amendments**
- Revise performance-level descriptors and proficiency ratings: **Hawaii**
- Include scores of students using calculators in the state mathematics assessments: **California**
- Revise the manner in which reading and writing scores are combined for AYP decisions: **Rhode Island** and **Wyoming**

**AYP Model**
- Clarify language regarding many elements in the state’s accountability workbook: **Alaska, Hawaii, Missouri, Texas, and Utah**
Adequate Yearly Progress (AYP) Application

- Restart AYP applications: Idaho
- Modify or clarify the manner in which AYP is calculated generally: Arizona, California, Hawaii, Massachusetts, Missouri, New Mexico, North Carolina, Oregon, and South Dakota
- Modify the manner in which new schools and reorganized schools or districts are defined: Arkansas, Connecticut, Missouri, New Jersey, New Mexico, North Carolina, North Dakota, Pennsylvania, Utah, and West Virginia
- Modify AYP determinations for schools not covered by state assessments and for small schools: Arizona, California, Colorado, Idaho, Missouri, Montana, New Mexico, North Carolina, Pennsylvania, and Wyoming
- Modify how school districts (or LEAs—local educational agencies) are identified for improvement: Oklahoma, Rhode Island, and South Dakota
- Modify or clarify the definition of Full Academic Year (FAY): Georgia, Hawaii, Louisiana, Missouri, Nebraska, Oklahoma, South Dakota, Utah, and Wyoming
- Base school and district identification for improvement on missing AYP targets in the same content area by the same student group for two consecutive school years: Oklahoma, Pennsylvania, Utah, and Washington
- Continue use of equi-percentile adjustments for some or all AYP determinations while transitioning to new content and achievement standards: Georgia and New Hampshire
- Exempt schools and districts with a break in Title I funds from improvement activities or exempt schools and districts not receiving Title I funds from the sanctions under section 1116: California and Missouri

Adequate Yearly Progress (AYP) Indicators

- Use indexing for AYP determinations: New Hampshire, New Jersey, and New Mexico
- Modify or clarify the manner in which participation rates are calculated: Colorado, Florida, Hawaii, Nebraska, New Jersey, Oklahoma, and Wyoming
- Modify Other Academic Indicators: District of Columbia, Iowa, New York, and Oregon
- Modify graduation rate targets, how graduation rates are calculated, use an interim graduation rate, or include students graduating past their scheduled graduation time: Alaska, Colorado, District of Columbia, Georgia, Iowa, Maryland, Massachusetts, Michigan, Nebraska, New Hampshire, New Mexico, North Carolina, Rhode Island, South Dakota, Tennessee, and Utah
- Permit the state to use or continue to use a growth model in AYP determinations or use confidence intervals in growth model AYP determinations: Delaware, Hawaii, Iowa, Louisiana, Missouri, North Dakota, Ohio, Oklahoma, Pennsylvania, and Utah

Annual Measurable Objectives and Intermediate Goals

- Modify starting points, AMOs, and IGs: Georgia, Idaho, Maine, Mississippi, New Hampshire, New Jersey, North Carolina, Oklahoma, Rhode Island, and Wyoming

Strategies to Enhance Reliability

- Modify the manner in which uniform averaging is used or discontinue use of two-year averaging in AYP calculations: Oregon, South Dakota, and Virginia
- Modify minimum n’s: District of Columbia, Massachusetts, Missouri, Montana, New Jersey, Ohio, Oklahoma, Oregon, South Carolina, Washington, and Wisconsin
Safe Harbor
- Modify the manner in which safe harbor is calculated: Colorado, Kansas, Missouri, New Hampshire, New Mexico, Texas, Utah, and Wisconsin

Inclusion of All Students

Inclusion of Students with Disabilities and AYP
- One-year extension of the flexibility to use “modified” achievement standards for certain SWDs:
  - OPTION TWO: Kansas, North Dakota, Oklahoma, and Texas
  - OPTION THREE: Colorado, Maryland, and Massachusetts
- Discontinue using “modified” academic standards: Alaska and New Hampshire
- Include, or discontinue including, exited SWDs in that student group when making AYP determinations: Alaska, Colorado, Connecticut, Hawaii, Indiana, Iowa, Kansas, Maryland, Missouri, Nevada, New Mexico, Rhode Island, South Dakota, Wisconsin, and Wyoming
- Modify the definition of FAY for the SWDs student group: Georgia
- Modify or clarify the manner in which the achievement of SWDs is included in AYP determinations: Hawaii, Louisiana, Maine, and Missouri
- Exceed the 1% cap for students with significant cognitive disabilities who “meet” or “exceed” proficient on alternate assessments: Massachusetts, Montana, and West Virginia
- Count special education diplomas as regular diplomas for AYP purposes: Georgia

Inclusion of English Language Learners in Assessments and AYP
- Modify or clarify the definition of LEP students: Alabama, Colorado, Connecticut, District of Columbia, and New Mexico
- Delay the time before which LEP students are required to participate in state assessments: Pennsylvania, South Carolina, Washington, and Wisconsin
- Modify or clarify the manner in which the achievement of LEP students is included in AYP determinations or in English language proficiency measures: California, Hawaii, Kentucky, Maine, Massachusetts, Missouri, Nebraska, New Hampshire, Virginia, and West Virginia
- Include, or discontinue including, exited LEP students in that student group when making AYP determinations and other flexibility options: Arkansas, Connecticut, Iowa, Kansas, Maine, Missouri, Nebraska, New Hampshire, and Utah
- Establish a “safe harbor” review for LEP students based on their progress on English language proficiency assessments: Georgia

AYP Consequences and Reporting

Public School Choice and Supplemental Educational Services
- Reverse the order of public school choice and supplemental educational services (SES) or permit schools identified for improvement to offer both school choice and SES in year one: Pennsylvania, Utah, and Virginia
• Apply consequences to the specific student groups not making AYP or allow for differentiated consequences: Utah
• Permit states to differentiate in the application of consequences to schools and districts not making AYP based on the degree to which targets are missed: Pennsylvania
• Clarify teacher qualifications: South Dakota

Reporting
• Delay or otherwise modify the release of school reports due to the implementation of new assessments: Missouri, Montana, New Jersey, North Carolina, and Texas

Appeals to Identification for Improvement Determinations
• Permit the state to consider supplemental data or otherwise clarify processes when schools or districts identified for improvement appeal that determination: Rhode Island and Utah
Chapter 1: Introduction

With the support of CCSSO’s Accountability Systems and Reporting State Collaborative on Assessment and Student Standards (ASR SCASS), the approval and amendment process for states’ educational accountability plans under the No Child Left Behind Act of 2001 (NCLB) has been chronicled in a series of annual papers since the initial submittal of these plans in January 2003. Last year, four years after the initial paper, it was assumed that 2007 would be a relatively quiet period for amendments. In the fall of 2005 and again in the fall of 2006, most states were busy implementing assessments in new grades, undergoing peer review of their standards and assessment systems, and adding science and alternate assessments based on modified achievement standards. Conversations about reauthorization of the Elementary and Secondary Education Act of 19652 had begun, although that abated considerably during late 2007. There seemed little need to make other major adjustments in accountability plans. Not only was that assumption quite wrong then, it proved to be wrong again this year. The range and total number of amendments requested remains at a higher-than-anticipated level.

Between 2003 and 2006, nearly every state submitted at least one request to amend its educational accountability workbook.3 While 47 states sought amendments in 2003-04, only about 20 submitted requests in 2004-05. In 2006, 47 states again submitted formal requests, 44 did so in 2007, and 49 states and the District of Columbia this year. In 2005-06, the accountability workbook amendments involved three surprises in terms of process. The first was the number of state amendment proposals. The second was even more surprising: a number of requests for changes previously approved for other states were denied. The rules, especially related to minimum n, definitely changed that year. The third was somewhat of a mid-year surprise when ED also solicited separate workbook amendments for states interested in reversing the order of school choice and supplemental educational services (SES) sanctions under NCLB (Spellings, 2006, May 15). Even though several states had already included these as 2005-06 amendments (and others did so last year), the Secretary’s announcement set up a separate application process as well as participation conditions. How the process works and how decisions are made remains unknown. (More details on states and districts approved are provided in Chapter 2.)

Last year, there were two more surprises with respect to the educational accountability workbook amendment process. The first was the manner in which ED handled states’ requests. The Department began by moving up the date for submittal of requests to February 15, 2007. Then, ED announced that it would include in its decision letters both amendment requests that were approved and those that were not (often with accompanying rationale). This was in sharp contrast to previous years in which responses—to the extent that they were known beyond an individual state—were not always quick or completely documented in writing, especially denials. By the end of August 2007, ED had responded to every state’s amendment request and posted a decision letter at the department’s website. States

---

2 NCLB is not a new law: it is simply the most recent reauthorization of the Elementary and Secondary Education Act of 1965 (ESEA). The most immediate prior reauthorization was known as the Improving America’s Schools Act of 1994. The first major section of ESEA, which includes the federal requirements for standards, assessments, and accountability, is referred to as Title I.

3 An NCLB requirement is that each state must develop an educational accountability system and a plan for the implementation and operation of that system. ED required that states describe the elements of their plans in a "workbook" format provided by the department. In this paper, references made to workbook amendments and plan amendments have the same meaning.
reported that this new policy helped them to better understand the status of their complete sets of amendment requests. The greater transparency was welcomed by all interested parties.

The second surprise last year was the attention ED gave to states using indexes and those that sought to add indexing applications in their AYP calculations. In the case of the former, the department invited states approved to use indexing (see Chapter 2 for more information) to a one-day conference in September 2007 to analyze “the impact of performance indexes on AYP determinations to ensure that they meet accountability” (Briggs, 2007, June 27). Continuing a course first observed last year, states seeking to add the use of indices in 2008 found themselves in concerted discussions with ED representatives.4

Undoubtedly, this year’s biggest surprise affecting states’ educational accountability system designs occurred after the accountability workbook amendment requests had been submitted for 2007-08. This surprise came in the form of a notice of proposed rulemaking (NPRM) issued by Secretary Spellings on April 23. The NPRM, fast-tracked by ED, with final regulations slated to go into effect on November 1, 2008, would impact a myriad of NCLB accountability provisions including, for example, explanations and peer review of how minimum group size and other components of AYP definitions (such as confidence intervals and indexes) ensure statistically reliable information; inclusion of NAEP data on state and local report cards; graduation rate calculations; criteria for growth models; increased state monitoring of supplemental educational services (SES) and stricter provisions governing parent notification for SES and public school choice. The measures were estimated by ED to cost states $22M to implement, an estimate most experts believe to be considerably understated. More detailed information is provided elsewhere in this chapter.

The second surprise this year, albeit not of the magnitude of the latest NPRM, came in the form of the number of states again submitting accountability workbook amendments and the extent of the requested changes. The number of the states, 49 and the District of Columbia, was exceeded only in 2003, when every State and the District was required to submit their first accountability workbook. And, the submittals and amendment proposals were also unanticipated given Assistant Secretary Kerri Briggs November 2007 letter to chief state school officers in which she stated that the department did not expect major changes this year given pending reauthorization. However, in addition to numerous minor “tweaking” of their accountability workbooks, the states submitted almost 275 substantive changes; 20% related to the manner in which AYP is applied; 20% related to the indicators used in making AYP decisions; 20% related to inclusion of SWDs and LEP students exclusive of modified achievement standards; 10% related to reliability measures; and 30% to other assessment and accountability issues.

This sixth installment of the Council of Chief State School Officers’ (CCSSO) papers reporting on annual changes to states’ educational accountability systems continues to document a public policy path that has become easier to track than in previous years. The structure of this latest paper parallels that of its predecessors. However, inconsistencies in ED’s decisions across states and over time also continue to underscore a fundamental principle reported in previous papers. Prior amendment decisions do not always represent trustworthy precedents for the future.

4Proposed regulations issued April 23, 2008, would codify a requirement that states modify and submit to ED for technical assistance and peer review their accountability workbooks to explain how the state’s current or proposed minimum n meets a new set of criteria and how other elements of the state’s AYP definition such as confidence intervals, performance indexes, and uniform averaging affect the statistical reliability of accountability determinations.
As with the previous five papers, this paper is based primarily on information provided voluntarily by states to CCSSO and subsequent decision letters posted by ED. States also submitted various types of information to support this work, including copies of their amendment request documents, emails, and reports of telephone conversations. CCSSO and the author are especially grateful to staff in state departments of education who have continued to make available the information essential to this public documentation of the amendment process and who so quickly responded to requests for more information or clarification of information.

In the remainder of this first chapter, the major issues that formed the 2007-08 accountability policy context are described. States’ amendment requests and ED’s responses to these requests are described in Chapter 2. Observations and conclusions are provided in Chapter 3, along with some speculation about what may lie ahead.5

The 2007-08 Policy Context

Accountability and accountability-related topics maintained their prominent position in the elementary and secondary educational policy spotlight during the 2007-08 school year. The year began with continued attention given to NCLB’s reauthorization. That attention, however, waned as it became clearer that reauthorization was not likely to receive serious consideration in the Congress until the next round of presidential elections concludes later this year. The year also began with announcements by ED of additional flexibility but appears to have quickly moved back to increased attention to regulation.

In terms of the former, growth models were opened to all interested states whose assessment systems were fully approved; a differentiated accountability pilot program was announced to create “a more nuanced system of distinguishing between schools in need of dramatic intervention, and those that are closer to meeting goals;” and the Secretary’s pilot program for school districts interested in reversing the order of public school choice and supplemental educational services was expanded. However, the April 23 NPRM proposing many new regulations pertaining to states’ educational accountability systems seemed to quickly shift the emphasis away from flexibility to enforcement.

As in prior years, a number of major themes emerged this year while others carried forward from last year, each in its own way seeming to both influence accountability workbook amendments sought by states for 2007-08, and ED’s decisions on these requests.

The 2007-08 themes include the following:

- Reauthorization Efforts Wane
- NCLB Litigation Challenges Continue
- More Reports on NCLB Effectiveness and Implementation
- English Language Proficiency Assessments Come in for Close Review

---

5 CCSSO’s Profiles of State Accountability Systems website tracks these systems nationally. With a focus on accountability under the No Child Left Behind Act, this site profiles each state’s accountability system and provides annual information on statewide standards, assessment, achievement, results, goals, reporting, and measurement systems. The site allows for the data to be viewed in multiple formats. One can go directly to a state to learn about its accountability system in a given year. One can also look at a particular element of statewide accountability across multiple states. While this site is password protected, interested parties can contact CCSSO for access. The direct link is [http://www.ccsso.org/projects/Accountability_Systems](http://www.ccsso.org/projects/Accountability_Systems)
Additional Flexibility—Growth Models, Differentiated Accountability, and Reversing Choice and SES
Latest NPRM Would Tighten and Require Additional State Justifications of Accountability Systems

The discussion of these themes that follows should be considered in light of the continuing and sometimes heated public and political attention to NCLB as reflected in communications from states, the press, related studies and reports, and papers presented at regional and national conferences.

Reauthorization Efforts Wane

Although reauthorization has been among the most frequently discussed and reported NCLB topics for more than two years, which changed as 2008 approached. The considerable attention given this subject earlier clearly lessened as it became more apparent that nothing was likely to happen in this vein until sometime into 2009. Nevertheless, debate and speculation surrounding reauthorization certainly remains among the key themes making up the 2007-08 policy context in which states have continued to move forward various proposals to amend their NCLB educational accountability workbooks.

The Status of Reauthorization—A Capsule

David Hoff, writing in Education Week late last year, summarized the fading prospects for reauthorization. He noted that Congress’ agenda was filled with other tasks and that “efforts to revise the law are mired in backroom negotiations in both the House and Senate and show no signs of gaining the momentum necessary to ensure completion of the reauthorization in 2008” (Hoff, 2007, p. 1). Hoff also pointed out that without reauthorization soon, the current law could remain in effect for another three years (p. 23). He commented that, “It’s difficult to enact major bills during a [presidential] campaign season, particularly with a president nearing the end of his second term; the arrival of a new president can also delay the [reauthorization] schedule. With the change in the White House, it often takes a year or more to finish detailed bills . . . that have been left hanging since the previous administration” (p. 23).

In a subsequent article on the status of reauthorization, Hoff (2008, January 16, p. 16) observed that “the policymakers who control the future of the No Child Left Behind Act agree the law needs changing. But they don’t agree on how extensive those changes need to be or how to enact them.” According to Hoff, some key lawmakers believe that major changes are needed but the President and the Secretary of Education disagree. Both have “suggested that they favor less significant changes to the law—and they are willing to use administrative powers to achieve their ends,” Hoff concludes. The subsequent announcements of a differentiated accountability pilot program (March), permitting the reversal of public school choice and supplemental educational services (April), and the proposal of additional regulations (April) seem consistent with the notion of “less significant changes to the law.”

---

6 The Education Commission of the States maintains a single source to track “who’s saying what” regarding NCLB reauthorization proposals. The database can be found online at www.ecs.org

7 Olson and Hoff (2006), in an Education Week article aptly titled “Framing the Debate,” provide an excellent context for the reauthorization debate that remains timely. They suggest that “renewing the law . . . [will] touch on at least a half-dozen ideas for changes. These include proposals for national academic standards; accountability models that measure students’ growth; testing and achievement targets for special populations . . . ; revisions in the sanctions for schools that miss . . . [AYP]; greater accountability for high schools; and more attention to teacher quality.”
Asian American Legal Defense and Education Fund (AALDEF)
Following a first-ever national summit for Asian American advocates, the AALDEF issued a report addressing several aspects of NCLB the organization believes need to be revised in order to better address the needs of Asian American students. The report’s authors note that, “Contrary to stereotypes that cast Asian Americans as model students of academic achievement, many . . . are struggling, failing, and dropping out of schools that ignore their needs” (Asian American Students Don’t Benefit from No Child Left Behind Act—Major Reforms Needed, 2008, p. 1). The report’s major recommendations included the following:

- Provide targeted language services for Asian American ELLs, since nearly a quarter of all Asian Americans are ELLs. Among those between the ages 5 and 17, over one-half of Hmong Americans, 39% of Vietnamese Americans, and 34% of Bangladeshi Americans are ELLs.
- Provide states with more funds to translate school documents, hire interpreters, and conduct community education for immigrant families. Over 40% of Vietnamese, Korean, and Chinese households are linguistically isolated.
- Require every state to collect comprehensive student data that is disaggregated by ethnicity, native language, socioeconomic status, ELL status, and ELL program type. Without this information, the educational needs of individual groups are concealed and will remain unaddressed.

NCLB Litigation Challenges Continue
This is an issue that has surfaced from time to time and one that seemed to “pick up steam” in 2005-06. Now, it has resurfaced. Two years ago, legislative committees in Colorado, Maine, Minnesota, and Utah had bills pending or proposed dealing with NCLB funding or other NCLB issues; Connecticut and the National Education Association had filed law suits; and, even a few school districts threatened to forego NCLB funding or otherwise challenged ED on issues related to interpretation and application of several of the law’s provisions.

National Education Association Suit
In early January of this year, a federal appeals court three-judge panel ruled that school districts or states cannot be compelled to use their own resources to carry out NCLB mandates. The ruling, which stemmed from a suit brought in 2005 by the National Education Association on behalf of 10 of its affiliate school districts, over-turned a federal district judge in Detroit who had dismissed the suit in November 2005. In response to the court’s decision, Secretary Spellings sent a letter to chief state school officers on January 18, 2008, indicating that she strongly disagreed with the decision and cautioned, “No state or school district should regard the ruling as license to disregard NCLB’s requirements.” Spellings further stated that she was exploring “all legal remedies to overturn the decision.” According to the Secretary, NCLB is not an unfunded mandate but rather a compact between a state and the federal government that asks the state and its school districts, in exchange for receiving substantial federal dollars, to demonstrate results. This investment in our children is creating opportunities for every child in America to have access to a high quality education. Indeed, states and school districts are seeing student achievement rise, especially for poor and minority children. Data show that NCLB is working. In my view, the Sixth Circuit’s decision undermines the efforts we have made under NCLB to improve the education of our nation’s children, particularly those children most in need. If the decision stands, it would represent a fundamental shift in practice.

Then in early May, it was reported that the full federal court (U. S. Court of Appeals for the 6th Circuit located in Cincinnati) had agreed to a request from administration lawyers on behalf of Secretary
Spellings to re-examine the three-judge panel’s decision. The agreement has the effect of setting aside the panel’s decision. No date was set for the re-hearing which served to keep lawsuit open (Walsh, 2008, p. 8).

**State Initiatives**

Several states are continuing to pursue previously initiated litigation while others are either taking steps in this direction or considering the same. In late April, a federal circuit judge in New Haven, Connecticut, rejected the few remaining issues in a suit originally filed by Connecticut in 2005. Other claims had been dismissed by the same judge in 2006. In bringing its suit, the state sought increased flexibility for testing SWDs and LEP students and also asserted that the NCLB funding was insufficient to provide the testing accommodations needed by LEP students or to develop alternate assessments for some SWDs. Connecticut’s attorney general indicated that the state would appeal the judge’s decisions (Walsh, 2007, p. 8).

In June, Arizona filed a lawsuit against ED claiming “that test scores of LEP students who have less than three years of English instruction are ‘psychometrically unreliable’ because ‘they may reflect only a student’s language proficiency rather than the student’s academic proficiency’.” Arizona’s state superintendent asserts that ED “reneged on an oral agreement given in 2003 that allegedly allowed the state to authorize schools to exclude the test scores of LEP students within the first three years of English instruction for the purposes of AYP.” According to the lawsuit, Arizona charges ED with breaching this agreement last year when the department insisted on enforcement of the one-year delay for inclusion of LEP scores in AYP determinations. The suit characterizes the one-year exemption as “arbitrary and capricious and not based on any scientific research or study.” In court documents, the state superintendent notes that he “specifically asked the official representing [ED] whether there was a scientific basis in education research that supported an exemption of one year rather than three years and was told there was none” (Brownstein, 2008, September, p. 6).

Also in Arizona, the state’s house passed a bill in April to permit the state to opt out of NCLB in 2010 provided that the state makes up any net loss in federal funding that would result. Passage in the state’s senate and enactment there were seen as unlikely. In Minnesota, some members of the legislature have once again signaled that they will seek passage of legislation to forfeit as much as $250M in annual funding the state receives under NCLB. Chief proponents argue that NCLB has diminished local control of schools and “grown” the federal role in schools. In Virginia, the legislature approved a bill in March “that would direct the state board of education to decide whether to withdraw from participation in . . . [NCLB].” However, it seemed unlikely that the governor or the state board would go along (Zehr, 2008, March 19, p. 16).

**More Reports on NCLB Effectiveness and Implementation**

Since 2005, numerous reports and studies have been released regarding the effectiveness of NCLB. Examples of these have been included in each paper since. More studies were released this year and more are likely as reauthorization nears. While the results are again somewhat mixed in terms of judging the law’s accomplishments and shortcomings, almost all conclude that NCLB is in need of substantive changes.

As noted in last year’s paper, an overdue NCLB mandated report remains to be issued. In this case, the law includes a provision in section 6161 that, beginning with the 2004-05 school year, ED will

> Conduct a review to determine whether each State has made “adequate yearly progress,” or AYP with respect to each subgroup accounted for under Title I. . . . Further, the law requires the
department to evaluate whether each State has met its “annual measurable achievement objectives” (AMAOs) for the attainment of English proficiency by limited English proficient students under Title III (Hicks, 2007, p. 1).

Examples of NCLB reports and studies released this past year are summarized below.

**The New Title I: The Changing Landscape of Accountability**

Published by Thompson Publishing Group, Inc., this is an update of the March 2005 edition. The revision is intended to guide readers “through the shifting landscape of NCLB and its core elements . . . by taking a look at where we are, then look back to see how we got here and finish by looking ahead to likely changes in the upcoming reauthorization” (Cowan & Edwards, 2007, p. xi). The report provides thorough, extensive background on the key provisions of NCLB; ED’s stewardship of the law—including how the department has at times demonstrated limited flexibility in targeted areas and at other times “stiffening enforcement of certain priority elements” (p. xi) along with a reauthorization forecast.

**National Assessment of Title I (October 2007)**

Presented in this two-volume report, together with a summary of key findings, are findings from the congressionally mandated National Assessment of Title I. The report centers on the implementation and impact of NCLB. Volume I is focused on NCLB’s implementation and Volume II on an evaluation of supplemental reading programs for 3rd and 5th grade students. Key findings are reported for Title I participants and funding; student achievement; implementation of state assessment systems; accountability and support for school improvement; AYP; school improvement activities; school choice and SES; and teacher quality and professional development. The full report is available at [http://ies.ed.gov/ncee/pubs/20084012/](http://ies.ed.gov/ncee/pubs/20084012/)

**Standards-Based Reform and the Poverty Gap (November 2007)**

Published by the Brookings Institution, a Washington think tank, this compendium of research examines more than two decades of federal efforts directed at increasing educational standards and improving the education of poor children. The intent is to determine what lessons can be drawn from all of these efforts to help NCLB achieve its goals of improving teaching and increasing student achievement, especially that of students from low income families. As summarized by the Brookings Institution Press (see [http://www.brookings.edu/press/Books/2007/standardsbasedreformandthepovertygap.aspx](http://www.brookings.edu/press/Books/2007/standardsbasedreformandthepovertygap.aspx)):

> Armed with the latest data and up-to-date research syntheses, the authors show that standards-based reform has had some positive effects, particularly in the area of teacher quality. Moreover, some of the critics’ greatest fears have not been realized: for example, retention rates have not shot upward. Yet the overall pace of improvement has been slow, owing in part to poor implementation. Based on these findings, the contributors offer recommendations for the implementation and impending reauthorization of NCLB.

---

8 In summer 2008, the Center on Education Policy released a study aimed at determining whether student achievement has increased and whether achievement gaps have narrowed between various student groups since the passage of NCLB in early 2002. The study ([Has Student Achievement Increased Since 2002?](http://www.cep-dc.org/)) by Kober, Chudowsky, N., & Chudowsky, V., which also include comparisons between state test results and NAEP results is available on line at [http://www.cep-dc.org/](http://www.cep-dc.org/)
Implications for NCLB’s reauthorization are presented in Chapter 10 of the compendium. Recommendations include national testing and rethinking of student academic achievement standards.

**Democracy at Risk—The Need for a New Federal Policy in Education**

Coinciding with the 25th anniversary of the release of the landmark publication *A Nation at Risk*, this report from another education think tank concludes that, “The federal strategy of attempting to improve schools through mandates and sanctions cannot get us where we need to go” (Darling-Hammond & Wood, 2008, p. iii). Other conclusions and findings include the following:

- Although many reforms have come and gone since 1983, we have lacked a purposeful, strategic approach for developing and investing in the kind of education that addresses the needs of a democratic society. In contrast to countries that have spent the last 20 years building forward-looking educational systems that fund schools centrally and equally, build a top-flight teaching force, focus on 21st century learning needs, and develop the capacity for school improvement, the U. S. has focused on none of these critical elements of success for an extended period of time (pp. iii-iv).

- The promise of NCLB has been its focus on achievement for all students and its insistence that all students have a right to highly qualified teachers. However, the law has not provided the resources to achieve these goals, and it has not focused on the kind of higher-order thinking and performance skills needed in the 21st century. These include the abilities required by social and democratic life to apply knowledge to complex, novel problems, communicate and collaborate effectively, and find, manage, and analyze information (p. iv).

- Federal policy under NCLB has encouraged schools to focus on a narrow band of knowledge, exhibited in ways that are not applied to important tasks in the real world. This policy has slowed gains in achievement in math and largely stopped progress in reading, while failing to reduce large gaps in school resources and student performance that are the major reason for America’s low rankings internationally (p. iv).

- The federal strategy of attempting to improve schools through mandates and sanctions cannot get us where we need to go. The successful new schools that have been created by many local reformers have been launched by educators and community members who together confronted old constraints; developed new curriculum, teaching, and assessment strategies; redesigned school organizations; and created learning communities that could drive ongoing improvement. These kinds of schools are constrained rather than enabled by top-down edicts and regulations. A new direction—grounded in what we know as a nation about innovation, learning, and powerful change—is needed to reclaim our leadership as a democracy and our children’s future in a land of opportunity (p. iv).

**English Language Proficiency Assessments Come in for Close Review**

Although all states have been required, under Title III of NCLB, to provide for the annual assessment of English language proficiency for all ELLs, the extent and adequacy of that assessment had not to receive much attention by ED until this past year. During 2007-08, there were four related, significant events. First, a major report released in late 2007 affirmed that all states now have the required annual proficiency assessments. Second, a March 2008 *Education Week* report confirmed that most states have begun to use the results of those assessments in making determinations of when ELLs no longer need specialized instruction (Zehr, 2008, March 26). Third, ED issued proposed interpretations of Title III requirements that would require states to further standardize the criteria they use to report how well ELLs
are doing and how they make decisions about when ELLs no longer require specialized services.\(^9\) Fourth, ED released its second biennial report to Congress on the implementation of Title III.

**All States Now Have Required Proficiency Assessments\(^10\)**

According to a research report issued in late 2007, all states now have in place and are using English language proficiency assessments—consistent with NCLB requirements—to measure the progress of ELLs in learning English. The research project, the collaborative effort of 32 experts on testing and the education of English language learners, and conducted under the leadership of Jamal Abedi (2007) at the University of California, Davis, looked at English language proficiency examinations that existed prior to the passage of NCLB and those that were created after passage of the law in 2002. According to the researchers, ED provided considerable support for states to develop their assessments through section 6112b Enhanced Assessment Grants. The assessments were developed mainly by four different consortia of states, although a few states proceeded on their own either in developing their assessments or augmenting “off-the-shelf” assessments from commercial vendors.

According to Abedi,

- The report is designed to provide information on the existing English language proficiency (ELP) assessments and discuss the national efforts in developing new ELP assessments based on the criteria required by NCLB Title III. The report also intends to share research, policy analyses, and technical documents from across the nation, which addresses critical Title III-related issues facing many state policymakers and educational leaders as well as educational consultants, publishers, and researchers (p. 6).
- The purpose of this report is to present facts about existing and newly developed assessments. We have no intention of evaluating the quality of existing and newly developed ELP assessments or criticizing any of these tests, whether they were developed prior to NCLB or after the law was implemented. Many of the existing ELP assessments have provided valid assessments for ELL students in the past. . . . We present some evidence on the content and psychometric characteristics of the existing ELP assessments since there were substantial data for these assessments. For the newly developed tests, however, we have not had as much data on the items to examine and discuss and must wait for the data from future test administrations before examining critical characteristics of these tests (p. 8).

**Role of ELP Assessment Results as Determinants for Specialized Instruction**

Perhaps fore-shadowing ED’s intention to seek greater standardization among states with respect to determinations of when students are ready to “exit” English language instruction programs, Mary Ann Zehr (2008, March 26), writing in *Education Week*, reported in late March on states’ efforts to develop guidelines in this area. According to Zehr (p. 8), “States vary widely in how prescriptive they are in the use of those test scores [English language proficiency assessments], but most seem to be taking steps toward standardizing the process.” Zehr conducted interviews with numerous state and LEA officials across the nation as well as prominent experts in teaching and assessing English language learners. Among her findings:

---

\(^9\) In late June 2008, ED sent a communiqué to State Title I Directors informing them that administration of the Title III state formula grant program (which includes the ELP assessments) would be moved in fall 2008 from the Office of English Language Acquisition to the Office of Elementary and Secondary Education’s Student Achievement and School Accountability Programs.

• Consistency is meaningful only if a state has followed the instructions under NCLB in developing tests that are aligned with state academic-content standards, and in basing instruction on those tests.
• Some district administrators contend that determinations such as reclassifying students as fluent in English or moving them out of specialized instruction should be made at the district, not state, level.
• The most widely used English-language-proficiency test is the ACCESS FOR ELLs test developed by the World-Class Instructional Design and Assessment [17-state member] consortium. . . . Most of those states have policies recommending or requiring that districts apply scores on the test to decisions on the ground. . . . [although] state officials are told they should let districts use a number of measures to decide whether ELLs no longer need specialized instruction.

ED Issues Proposed Interpretations of NCLB Title III Requirements
In somewhat of an unusual move, ED issued a Notice of Proposed Interpretations on May 2, 2008, regarding several provisions of Title III of NCLB. In six years of implementing NCLB, the department had not previously issued interpretations of the law, apparently choosing instead to issue policy letters from time-to-time. The Notice essentially tells “states they must use a consistent yardstick in determining when a child is fluent in English and when the child no longer needs special ELL services.” Further, “states must . . . standardize the criteria they use to report how well such students are learning English” (Zehr, 2008, May 14, p. 1). The final interpretations had not been issued as this paper was completed.

The most immediate impact of the proposed interpretations would be to curtail the flexibility that many states have given to school districts in these matters, especially with respect to determining when students are ready to exit special services for limited English language learners. Such an action would be contrary to what state and local education officials believe according to Zehr’s recent survey findings noted above. In a follow-up on the proposed interpretations after the 30-day public comment period ended on June 14, 2008, Zehr (2008, June 11, p. 17) further observed that “education officials in several states with large English-language-learner populations are bristling at a proposal . . . that they say would curb their flexibility in deciding when children are fluent in English and if they still need special services for ELLs.”

Another area of contention that emerged is that of “banking” proficient scores on English language acquisition assessments. The proposed interpretations would require states to provide an annual assessment of English language proficiency that measures the oral language (speaking and reading), reading, and writing skills of all public school LEP students. ED’s position is that the law clearly requires that LEP students be assessed in all domains on an annual basis and, thus, prohibits the “banking” of proficient scores in a particular domain until such time as the student is proficient in all domains and exits the LEP student group (Notice of Proposed Interpretations, 2008, May 2).11

Second Biennial Report to Congress on the Implementation of Title III
In late June 2008, ED sent to Congress its second biennial report on the implementation of Title III. The report covered the school years 2004-05 and 2005-06.12 Key findings included (pp. vii-viii):

11 ED has permitted states to bank test scores under Title I under certain conditions.
12 The report is available at http://www.ed.gov/about/offices/list/oela/title3biennial0406.pdf
The number of K-12 students in the United States who are identified as being LEP has grown by nearly 650,000 in the past three years and is now approximately 4,985,000.13

LEP students speak more than 400 different languages, including languages from outside the United States (e.g., Asian, European, and African languages) and inside the United States (American Indian languages). Nearly 80 percent of LEP students speak Spanish; another 5 percent speak Asian languages.

The National Assessment of Educational Progress (NAEP) indicates that the national sample of LEP students have improved their performance in the content areas of mathematics and reading or language arts. For mathematics, the percentage of fourth- and eighth-grade LEP students who scored at or above the basic level was higher in school year 2007 than in any previous year. For reading, the percentage of fourth-grade LEP students who scored at or above the basic level was higher in 2007 than in 2005; eighth-grade LEP students showed no increases from 2005.

During school year 2005–2006, officials in 31 states and the District of Columbia tracked the continuing education progress of more than 312,000 students who were formerly classified as LEP; 86 percent of these students scored at the proficient level or above in mathematics and 99 percent scored at the proficient level or above in reading or language arts.

Additional Flexibility—Growth Models, Differentiated Sanctions, and Reversing Choice and SES

In late 2007 and early 2008, Secretary Spellings signaled that she would be traveling extensively across the country over the first part of 2008 seeking ideas for NCLB reauthorization and states’ ideas for additional flexibility in the implementation of NCLB. The three initiatives described below appear to have resulted from the Secretary’s interest in providing additional flexibility around NCLB’s implementation. However, these initiatives stand in somewhat of a sharp contrast to the prescriptive set of proposed new regulations issued by the Secretary on April 23, 2008. Those proposed regulations are described in the following section of this chapter.

**Growth Models**

Soon after NCLB’s enactment in 2002, a few states began advocating for the use of growth models in making AYP decisions. These states argued, unsuccessfully then, that measuring the progress of individual students over time was a more valid and reliable approach than making AYP decisions based on status; for example, one group of third graders compared to a succeeding group of third graders. Then, in late 2005, Secretary Spellings (Spellings, 2005, November 18) introduced a growth model pilot program to allow up to 10 states to implement “high quality” growth models in their adequate yearly progress (AYP) calculations for the 2005-06 school year. Following peer reviews, two states (North Carolina and Tennessee) were approved in May 2006 to implement growth models for the 2005-06 school year. Three more states (Arkansas, Delaware, and Florida) were approved in November 2006 for the 2006-07 school year; two more (Iowa and Ohio) were approved in May 2007 for the 2006-07 school year; and two more (Alaska and Arizona) were approved in July 2007 for the 2006-07 school year.14 In a few instances, approval was conditioned on a state receiving full approval of its assessment system or establishing a uniform minimum n.

---

13 The Census Bureau has projected that the rise of minority groups, particularly Latinos and African Americans, will help propel minority groups in the aggregate to be the new majority by 2042.

14 For further information and resources on state accountability systems, including growth models, see http://www.ccsso.org/projects/accountability_systems/
On December 7, 2007, Secretary Spellings announced that she was opening the growth model option to all eligible states, defined as those with approved statewide assessment systems. A February 1, 2008, deadline was set for applications. Applications were to be subject to a peer review process following ED’s determination that they met the seven core principles set earlier by the department for growth models and that the applying states were making progress in specified areas. Additionally, proposed new regulations issued on April 23, 2008, would codify non-regulatory criteria that have been used by ED to make approval decisions related to growth model proposals.

In June, ED announced the approval of two additional states (Michigan and Missouri) to use growth models in their AYP determinations. At the same time, however, ED appears to have added yet another previously unannounced criterion with respect to growth model reviews. Specifically, in the case of both Minnesota and Pennsylvania, ED’s primary basis for rejecting their growth model proposals appears to have been because each state uses indexing as an AYP “screen.” Assistant Secretary Kerri Briggs’ decision letter (2008, June 10) to Pennsylvania is illustrative:

The heart of the peers’ concerns relates to the interaction of Pennsylvania’s existing performance index with the growth model. The Department has concerns about the appropriateness of allowing a state to include both a performance index and a growth model in its accountability system. I will task the Department’s recently announced National Technical Advisory Council to review this issue and provide feedback on if, and when, it may be reasonable for a state to include both components in its accountability system. We will provide additional guidance on this question this summer so that Pennsylvania may take this guidance into account when developing a revised proposal. [The guidance had not been issued as this paper was completed.]

**Differentiated Sanctions**

A second area of flexibility put forth over the past year by ED was that of a new pilot program aimed at a differentiated approach to the application of NCLB sanctions in states’ accountability systems. Announced by the Secretary in mid-March, up to 10 states were to be permitted flexibility with respect to how resources and interventions could be targeted to the lowest-performing schools. Again, as with growth models, states interested in participating in the pilot program must have approved statewide assessment systems. ED also announced that it would use a separate application and peer review process to make its approval decisions. According to a fact sheet released by ED in May 2008, 17 states (Alaska, Arkansas, Florida, Georgia, Illinois, Indiana, Louisiana, Maryland, New Jersey, New York, North Dakota, Ohio, Oklahoma, Pennsylvania, South Carolina, Tennessee, and Virginia) submitted applications to participate in the pilot program. At the same time, the Department announced the appointment of a 15-member peer review panel that would meet in June.

In peer review guidance issued on April 3, 2008, ED set forth three unique terms to “describe the complex ways” states might propose “to differentiate accountability.” The terms “phase of improvement,” “category of improvement,” and “capacity cap,” all refer to sanctions that are applied once a school or district is identified for improvement. Although ED refers to this new pilot as a “differentiated accountability model,” it is essentially more a means of labeling, describing, and carrying out various sanctions that must be applied under NCLB (see sec. 1116 of the law) when schools and districts do not make AYP targets over a period of two or more years. ED characterizes the model as “a more nuanced way of distinguishing

15 On July 1, 2008, ED announced that six states had been approved to participate in the pilot program—Florida, Georgia, Illinois, Indiana, Maryland, and Ohio. Additional information is available at www.ed.gov/admins/lead/account/differentiatedaccountability/
between schools and districts in need of intensive intervention and those that are closer to meeting their goals." The pilot does not, as implied, permit states to implement any other educational accountability system besides that prescribed under NCLB. States cannot, for example, use a compensatory model instead of a status model for AYP determinations. Following the June 13-14, 2008, peer review, ED released a paper on “cross-cutting issues” identified during the peers’ deliberations that “helped them recognize the strengths, shortcomings, and pitfalls of differentiated accountability models.” The paper is available at http://www.ed.gov/admins/lead/account.differentiatedaccountability/index.html

ED included the provisions for the “differentiated accountability model” in its April 23, 2008, notice of proposed rulemaking.

**Reverse Public School Choice and Supplemental Educational Services Sanctions**

Somewhat of a surprise was the third area of additional flexibility announced by ED in early June. After years of consistently denying states’ requests to be able to reverse the order of public school choice and the provision of supplemental educational services (SES)—with the exception of a pilot program begun in 2005-06 for a limited number of school districts—the department reversed course saying that all states could apply for this flexibility. However, ED will still consider it a pilot program and maintain the separate approval process it has put in place outside of the accountability workbook amendment process. The caveat again is that only states with approved statewide assessment systems (34 states as of June 2008) would be eligible. A primary goal of the department underlying this decision is the desire to increase student participation in SES. The April 23, 2008, proposed regulations include a number of very specific, new procedures participating states will have to implement, many of which carry appreciable fiscal implications for state education agencies. These include (Mesecar, 2008, June 4):

- Increase student participation in SES.
- Maintain a comprehensive list of approved SES providers.
- Ensure that there are at least two providers available in each participating school district.
- Offer continuous enrollment in SES or multiple SES enrollment periods throughout the school year.
- Ensure that school districts provide fair and equitable treatment of non-school district providers by giving access to school facilities at reasonable prices and dividing space among providers in a fair manner.
- Allow providers to market their services and work with community and business partners to reach parents.
- Submit various additional reports to ED.

**Latest NPRM Would Tighten and Require Additional State Justifications of Accountability Systems**

Stating that she was proposing “new policy tools that will give families lifelines—and empower educators to create dramatic improvement,” ED Secretary Margaret Spellings announced on April 22, 2008, the promulgation of an extensive set of far-reaching proposed new regulations governing NCLB. The Secretary characterized the proposed rules as “policy bulldozers” intended to “tear down barriers to [educational] reform” and help to put out “the ‘raging fires’ in our lowest performing schools” (Spellings, 2008, April 22). The rules were “fast-tracked” for final issuance by November 1, 2008, to be effective one month later.
The notice of proposed rulemaking (NPRM)\(^{16}\) was published in the *Federal Register* on April 23, 2008, and, if codified, would affect at least nine broad areas of states’ assessment and educational accountability systems—many carrying considerable additional cost implications for states.\(^{17}\) In its background statement related to the proposed rules, ED concludes, “These proposed regulations build on the advancements of State accountability systems, while incorporating key feedback from the field on an even clearer vision of what it takes to educate each and every one of our Nation’s schoolchildren” (pp. 22020-1). However, neither the sources nor the substance of “key feedback” were provided.

The proposed regulations address the following major areas:

1. Clarifying assessments and multiple measures— Attempts to clarify what multiple measures might be and what types of questions are illustrative of those.
2. Establishing a National Technical Advisory Council— Intended to “advise the Secretary on key technical issues related to State standards, assessments, and accountability systems that are part of State plans.”
3. Validating minimum student group size including re-justification of measures states use to ensure statistical reliability of their accountability determinations— No later than six months following the effective date of the new regulations, states would be required to submit for another peer review revised accountability workbooks (all states currently have ED-approved accountability workbooks) “to explain how the State’s current or proposed minimum group size” meets new regulations related to establishing minimum n’s. According to comments in the NPRM (p. 22023), “States that propose large minimum subgroup sizes and include other components in their AYP definitions that result in the exclusion of large numbers of students or student subgroups would be subject to close scrutiny.” Along with minimum group size justifications, states will also be required to include in their accountability workbook re-submissions explanations of how “the State’s AYP definition (such as the use of confidence intervals, indexes, and uniform averaging [plus the definition of FAY]) . . . affect the statistical reliability of accountability determinations as well as impact the inclusion of all students and student subgroups in those determinations.”
4. Including NAEP results on state and school district reports— Intended to provide parents and the public “with a tool to compare how students in a State are performing on the NAEP with performance on State assessments.”
5. Defining, calculating, and reporting graduation rates— Proposes numerous changes, including use of the National Governors Association definition for the purpose of establishing a “uniform and accurate method of calculating graduation rates;” requiring states to strengthen their graduation rates goals and measures of continuous and substantial improvement; and disaggregation of rates and inclusion of those in AYP determinations.\(^{18}\)


\(^{17}\) The NPRM reported the estimated costs to be approximately $22M. Most experts believe this figure to be substantially understated.

\(^{18}\) An April 4, 2008, report from ED’s Office of Inspector General concludes that, “OESE [Office of Elementary and Secondary Education] could have provided better oversight to ensure that graduation and dropout rates submitted by states in the Consolidated Performance Reports were supported by reliable data. Specifically, we determined that the Department did not put enough emphasis on data reliability and comparability across states.” The report further concluded, “If the Department had been more assertive in requiring states to implement a longitudinal student tracking system shortly after the enactment of NCLB, all states could now have four years of student data. Instead, less than a quarter of the states are using a system that complies with the requirements of the law” (West, 2008, April 4).
6. Including growth models in AYP—codifies criteria ED has published related to peer reviews of growth models and additional criteria emanating from peer reviews. Although not specifically provided for in the NPRM, the regulations would apparently continue to force states and districts to forego the use of confidence intervals and indexing in any AYP decisions, both prohibitions ED has applied in its approval decisions. As provided for on page 22027, “States could permit its LEAs and schools to make AYP by meeting (1) the State’s proficiency targets, (2) growth targets, or (3) the ‘safe harbor’ provision.”

7. Prohibiting AYP determinations based on same AYP targets for two consecutive years—although NCLB is silent on the matter, ED first insisted that schools and districts missing any of the potential 37 AYP targets for two consecutive years would be identified as in need of improvement. However, that decision was quickly amended to any target in the same subject area two years in a row. The proposed regulation would codify that position, a position that many states have continued to challenge since 2003, arguing that the intent of the law is to focus improvement efforts and resources on the specific student group(s) failing to make various AYP targets for two or more consecutive years.

8. Clarifying public school choice and the provision of supplemental educational services (SES)—includes a number of new requirements for states, particularly with respect to SES intended to boost student participation in both of these. School districts would have to coordinate with community groups to help advise parents of their options; parents would have to be notified not less than 14 days prior to the beginning of school each year; SES open enrollment would have to be available throughout the school year, and states would have expanded responsibilities related to approval and evaluation of SES providers, monitoring SES implementation, and giving providers greater access to school facilities. The new requirements hold the potential for significant cost increases for state education agencies related to monitoring, noticing, and reporting.

9. Systematizing differentiated accountability determinations (differentiated applications of sanctions under the law)—codifies the “differentiated accountability model” announced in early April (see above for discussion).

Andrew Brownstein (2008, June), writing in the Title I Monitor, opined that the new regulations related to increasing graduation rates (thus reducing dropout rates) would effectively result in more schools missing NCLB AYP targets. According to a National Education Association director Brownstein interviewed, the regulations on graduation rate (which would have to be disaggregated and included in AYP determinations) would add nine additional categories to the current accountability cell matrix of 37 (p. 1). Brownstein concluded his analyses regarding the impact of the proposed rules on graduation rate determinations by quoting the same association executive who characterized the process as “reauthorization through regulations” and another who observed that they “change 20 aspects of current NCLB law and regulation . . . ” (p. 6). The Secretary is also reported to have stated at a press conference on April 22, following announcement of the NPRM, that “she is using her administrative powers because Congress and the Bush administration had been unable to agree on a reauthorization” (Hoff, 2008, April 30, p. 25).19

19 According to the NCLBonline produced by the Thompson Publishing Group, Inc., 382 comments were submitted to ED by the June 23, 2008, deadline (a large number of duplicative “boilerplate” comments generated by a single web site posting were not included in the count of comments). According to the NCLBonline posting, “The comments reflected the full range of perspectives . . . with civil rights advocates strongly supporting the administration’s proposals . . . and local agencies decrying the expanded administrative burden and confusion they claim will be
In a follow up article, Brownstein (2008, August) presented an overview of the responses received by ED regarding the NPRM. He noted that reactions “followed predictable lines, with civil rights advocates generally supporting tougher enforcement of the law and states and districts claiming that the department is making it tougher to comply with the complex requirements of No Child Left Behind” (p.1). However, Brownstein noted, “this time . . . the criticism has been sharper” (p. 1) with cost, timetables, and “legislation by regulation,” among the principal criticisms.

**The Next Chapter**

Educational accountability workbook amendments states requested during 2007-08 are reported in Chapter 2, together with ED’s decisions on these proposed amendments. As consistently noted in this series of papers, the information reported should be interpreted in light of two important considerations. First, each paper has relied on publicly available state requests for accountability amendments. Some states occasionally request amendments without making these requests public or sometimes submit additional requests (via email or verbal communications) that are not shared with CCSSO or posted by ED; thus, it is possible that some requests have been made that are not captured here.

Second, neither CCSSO nor the authors who have contributed to this series necessarily endorse any of the educational accountability strategies embedded in the described amendments. The purpose of these papers continues to be to chronicle accountability system plan amendment proposals, not to judge any state’s intent in making a given proposal, or whether the proposals are appropriate or technically sound. Neither is this paper intended to serve as an evaluation of ED’s responses or failure to respond as may be the case. No recommendation of sound, appropriate, or effective practice is implied or should be inferred.

caused by the proposals.” The comments are also available at [http://www.regulations.gov/search/index.jsp#](http://www.regulations.gov/search/index.jsp#) (enter “Title I” in the search option).
Chapter 2: 2007-08 Amendment Requests and ED’s Decisions

In this chapter, the educational accountability plan amendments that States have requested (or, in a few cases, were directed by ED to make) since the beginning of the 2007-08 school year are summarized as are ED’s responses to these requests. In spite of a caution by Elementary and Secondary Education Assistant Secretary Kerri Briggs that states limit their amendments, the number of requests this year was again surprising coming some six plus years after NCLB’s enactment (Briggs, 2007, November 19).

Fortunately, ED’s decision for all states requesting amendments this year are known as this paper is completed, reflecting the department’s decision to respond in writing to all states submitting accountability workbook amendment requests beginning in 2006-07. It must be noted, however, that although ED has made public its decisions on nearly all amendment requests, the response letters have not always reflected the position communicated to state officials during the negotiations preceding the issuance of the letter.

Source documents for material quoted directly from a state’s submittal of proposed accountability workbook amendments or from a U. S. Department of Education (ED) decision letter are not included in the list of references at the end of this paper. The former are available on request to individual states and the latter are posted at http://www.ed.gov/admins/lead/account/letters/index.html.

The chapter is organized into the following major categories:

- Standards and Assessments
- Adequate Yearly Progress (AYP) Model
- Inclusion of All Students in Accountability
- AYP Consequences and Reporting

Standards and Assessments

From early 2005 through the end of 2007, every state underwent a peer review of their final assessment systems under NCLB. These reviews involved the submission and often re-submission multiple times by each state of substantial amounts of descriptive narrative and documentation in relation to ED’s Standards and Assessments Peer Review Guidance (2004, April 24). According to David Hoff, writing in Education Week, ED had approved the assessment systems of 31 states by the end of March 2008; another four were making final changes requested by the Department; nine were making revisions that would not be completed until at least the upcoming school year; and six other states were developing agreements to be in compliance within the next three years (2008, April 2, pp. 1 and 24). Technical quality of the alternate assessments for students with significant cognitive disabilities has been a particularly challenging barrier to attaining full approval in most of these cases. Also, a few states have

20 The guidance was revised to include modified academic achievement standards and re-issued by ED on December 21, 2007.

21 In a letter to chief state school officers in early June 2008, Assistant Secretary Kerri Briggs stated that only 32 states had full approval or full approval with recommendations and that it would be unlikely that the remaining states would reach this status by the end of the current school year. As a result, at least four states have already entered into Compliance Agreements. An exception appears to be Texas where the Department is requiring the state to enter into a Memorandum of Understanding instead of a Compliance Agreement “to address all outstanding issues” with the Texas Assessment of Knowledge and Skills-Alternate (May 7, 2008).
recently brought online new assessments based on new and strengthened academic content standards. This has served to delay final approval as well.

ED has continued to conduct standards and assessment peer reviews since the end of last year. However, the focus has shifted to alternate assessments based on modified academic achievement standards and science assessments. States were required to implement the latter beginning in 2007-08, in at least one grade in each of the spans 3-5, 6-9, and 10-12. States are not required to develop alternate assessments based on modified achievement standards although many are doing so.

The second-most common accountability workbook amendment this year related to adding science assessments to states’ systems. States also asked for changes related to their overall assessment systems, expansion of student participation in Advanced Placement and International Baccalaureate exams in lieu of state assessments, use of score banking, recalculation using “best scores,” and other miscellaneous standards and assessments amendments.

There were no requests related to high school exit examinations or end-of-course examinations in connection with AYP determinations. According to a recent study by the Center for Education Policy (Zabala, et al, 2008), 26 states have current or planned exit examinations. By 2012, when all 26 states have implemented their exit exams, “approximately 74% of the nation’s public high school students will be affected” (p.1). Additionally, “more than 84% of students of color will live in such states” (p. 2). The report’s authors also observed a slow, but steady increase in the number of states using, or planning to use, end-of-course examinations—from two in 2002 to a projected 14 in 2015 (p. 2).

Changes to Assessment Systems

Science Assessments

With respect to changes to state assessment systems, most of the “action” centered on efforts to implement new science assessments in the grade spans 3-5, 6-9, and 10-12 or to otherwise add those assessments to educational accountability workbooks. Under NCLB, states are required to administer, beginning in 2007-08, annual science assessments in at least one grade in each of these three spans. In many instances, states had been administering science assessments prior to 2007-08. In late February 2008, Assistant Secretary Kerri Briggs sent information to chief state school officers regarding the process and criteria for peer review of these assessments (Briggs, 2008, February 28). Student achievement results on the science assessments are not required to be included in AYP determinations unless a state elects to use them in some manner.

Almost without exception, ED’s approval (when cited in decision letters) of states’ plans for the administration of their science assessments appears to have been routine; albeit with occasional caveats that final approval would be conditioned on peer review outcomes. One pattern of approvals at the high school level, however, merits additional exploration or comment here.

Most states seem to have developed comprehensive, cross-domain science tests at all grade levels assessed. Typically, these states assess up to four science domains and all students are expected to

---


23 Information on States that have made passing a high school exit examination a prerequisite to graduation and the attributes of those examinations can be found at http://www.hsee.umn.edu
take the same assessment at a designated grade within each of the three grade spans. However, this approach did not hold true with respect to ED’s approval of high school end-of-course science assessments for some states. For example, **Utah** students in grades 9 – 12 are required to take two science courses for graduation. They choose the required courses from earth systems, biology, physics, and chemistry. For each science course taken, the students take a CRT. The state then determines proficiencies—by course—for every science CRT given in each school each year to meet the NCLB requirements. In this approach, then, not all students take the same science assessment for NCLB purposes but all students do take at least two science CRTs while in high school albeit in a variety of domains.

**Virginia** administers end-of-course (EOC) assessments in a number of subject areas at the high school level. For high schools there, the state annually rolls together student scores for EOC assessments in earth science, biology, and chemistry to create a “pass rate” for science (that is “high school” but not grade-specific). The science “pass rate” is not used in AYP at the state level although a few school districts choose to use science scores as the OAI.

According to Virginia state officials, the same science EOC test is not required for all students at the high school level. Earth science, biology and chemistry each have an EOC test but not physics. Generally, earth science is taught at 9th grade, biology at 10th grade, and chemistry in 11th grade, however the VDOE does not prescribe a sequence of courses. An alternate assessment in science is provided at the 11th grade for students with significant cognitive disabilities.

At least 16 states and the District of Columbia submitted accountability workbook amendments or updates in this area. **Alabama** will administer its new assessments in grades 5, 7, and 11. **Alaska** will administer its science assessments at grades 4, 8, and 10. In **Arizona**, the state will assess in grades 4, 8, and 10. At the high school level, 9th graders may take the science assessment and have their scores banked for reporting the following year (approved July 15, 2008). In the **District of Columbia**, science assessments will be administered in grades 5 and 8 and in biology at the high school level (approved July 28, 2008). **Maine** clarified that its high school science test will be the Maine High School Assessment science test (approved September 19, 2008). **Missouri**’s science assessments will be given in grades 5, 8, and 11. Student scores will be included in the data used as part of Missouri’s Annual Performance Report for the classification of public school districts there.

**North Carolina**, which will assess science in grades 5, 8, and high school, also plans to use student scores as part of its State Composite Calculation (but not for AYP). At the high school level, ED previously approved North Carolina’s use of its biology end-of-course assessment at the 10th grade (scores are banked for students testing prior to 10th grade). However, the state will change the science administration from grade 10 to grade 11 because “several high schools have different science-taking sequences.” Extending the point of administration would ensure that more students are included (approved June 6, 2008). **Hawaii** will administer its science assessments in grades 5, 7, and 11 (approved September 25, 2008). In **Illinois**, students will be assessed in biological and physical sciences in grades 4 and 7. **Kansas** will administer its science assessments in grades 4, 7, and 11. **Kentucky** and **Wisconsin** clarified their procedures for administering science assessments during 2007-08 and future years. **Michigan** will administer its assessments in grades 5, 8, and 11 (approved July 15, 2008).

**Maryland** indicated that students there will take the science assessments in grades 5, 8, and high school. At the latter level, the state received approval (June 17, 2008) to use its high school biology assessment because Maryland students must pass this test as a requirement to receive a state high school diploma. Maryland also received approval to substitute a score of 3-5 on the Advanced Placement (AP) Biology
examination or a score of 5-7 on the International Baccalaureate (IB) Biology SL or Biology HL for the high school biology assessment. The state had previously received approval to substitute AP and IB examinations for end-of-course assessments in English and algebra/data analysis.

In Pennsylvania, the state will administer its science assessments (measuring content and inquiry) in grades 4, 8, and 11 involving approximately 400,000 students (approved August 18, 2008). Utah, which has administered criterion-referenced tests (CRTs) in science for many years, developed a somewhat different approach. Students there in grades 4 through 8 take a grade-specific CRT since they are enrolled in grade-specific science courses. However, in grades 9-12, students are allowed to select from a variety of science courses. Therefore, they will take a science CRT in earth systems, biology, physics, or chemistry (approved June 17, 2008). Virginia will use its already administered science assessments in grades 3, 5, and 8, as well as end-of-course assessments at the high school level. Results will continue to be used in the state’s accountability system (but not in AYP decisions). On June 24, ED approved the plan but cautioned, as with other states, that this did not constitute “approval of Virginia’s science assessments and, in approving this amendment, the Department expresses no opinion on the sufficiency of Virginia’s science assessments.”

Other Assessment System Changes
Colorado revised its alternate assessment in reading for 2007-08 to “reflect both the level of assistance a student needs to answer a question and whether or not the student answers the question correctly.”
Illinois submitted two changes to its assessment system. The state updated its schedule of assessments to reflect current practice and removed references to the Illinois Measure of Annual Growth in English from its workbook. Illinois, which previously used this assessment with LEP students, will have them take the regular assessments with or without accommodations. According to a brief article in Education Week, Illinois was unable to convince ED that the test is comparable to its regular assessments and the state will develop another alternate for its 36,000 ELLs. While ED has also required Arkansas and Wisconsin to stop using their alternate assessments for this student group, it has approved similar tests in North Carolina and Virginia (Zehr, 2007, November 7, p. 4).

Missouri reported that it will change in 2008-09 from use of its Missouri Assessment Program tests to end-of-course assessments at the high school level. The state will provide three testing windows for the new assessments—fall, spring, and summer. In a June 25, 2008, decision letter to the state, ED stated that, “Since the assessments are not yet part of the Missouri standards and assessment system and evidence regarding these assessments has not been submitted for NCLB assessment peer review, this request may not be included in the accountability workbook at this time.” New Mexico, citing “fiscal constraints,” will drop its 9th grade assessments. For schools enrolling only 9th graders, the state will apply its “feeder-school methodology” to make accountability determinations (approved July 23, 2008).

North Carolina sought permission to substitute “different NCLB assessment requirements for certain innovative schools (New Schools Project, Learn and Earn Schools, as examples) that offer assessment alternatives that better align with the objectives of the innovative schools than do the current statewide assessments.” ED denied the request on June 6, 2008, stating, “If North Carolina wishes to pursue the use of substitute assessments, it must submit such assessments for peer review before they may be used in AYP determinations.” Pennsylvania’s governor and lawmakers have decided to postpone implementation of requirements for high school students there to pass a set of 10 subject-specific tests in order to graduate. Instead, participation in the tests will be optional for the time being.

West Virginia’s State Board of Education has adopted new academic content standards that will be effective in 2008-09. The state will also first administer new assessments aligned with those standards at
the same time. ED’s July 15, 2008, decision letter noted that the new standards and assessments would need to be peer reviewed adding, “in approving this amendment, the Department expresses no opinion on the sufficiency of either West Virginia’s new content standards or its new assessments.”

**Use of ACT, SAT, AP, or IB Exams as High School Accountability Assessments**

The use of college admissions tests, Advanced Placement (AP) tests, or International Baccalaureate (IB) examinations for AYP purposes was first detailed in the 2006 paper on accountability workbook amendments (Forte and Erpenbach, 2006, pp. 12-13). More recently, Achieve, Inc. also examined the use of college admissions and placement tests as statewide assessments under NCLB (Achieve, 2007).

The intent of the study was to “help inform the decisions states are making about high school assessments by providing greater insight into the world of college admissions and placement testing” (p. 2). The study concluded that “College admissions and placement tests vary considerably—and do not fully measure the knowledge and skills that are included in the ADP [American Diploma Project] benchmarks” (p. 2). The authors of the study offered four recommendations for K-12 policymakers (pp. 3-4):

1. Augment admissions tests when incorporating them into statewide testing systems.
2. Consider using end-of-course tests to tap higher-level content and skills and place students into college courses.
3. Modify existing high school tests to measure college readiness.
4. Use existing college placement tests for diagnostic purposes only.

In mid-March of this year, it was reported that legislation being advanced by Republican lawmakers in Kentucky would repeal the state’s current testing system in favor of new assessments including requiring the ACT in place of high school tests (Jacobson, 2008, p. 17). At the end of April, ED announced that it had given full approval to Maine’s use of an augmented SAT examination to test high school students against the state’s academic content standards at the 11th grade. Maine originally began using the SAT for this purpose in 2006, making adjustments in 2007, by adding augmented items in the mathematics portion of the test. Maine has consistently argued that use of the SAT supports its efforts to promote post-secondary education by having all students take a college admissions test such as SAT.

To date, only two states have sought to use AP or IB examinations in their assessment systems. In 2006, ED granted flexibility for one year permitting Maryland to use AP or IB exams in lieu of the end-of-course Maryland High School Assessments (HSA) for a small number of students. In 2007, ED extended this option to Virginia (July 16, 2007). Maryland originally argued that the “university-level work represented by AP and IB assessments exceeds the expectations of the Maryland High School Assessments and that these highly rigorous assessments in one of the related disciplines require a student to have fully mastered the content in a foundational course.” Under the one-year flexibility approval, students who score 3 or higher on the calculus, English language, and English literature AP exams or those who score 5 or higher on the mathematics studies, mathematics, and English 1A International Baccalaureate exams were considered proficient in high school AYP determinations for algebra/data analysis and English 2.

Last year, Maryland also received approval (August 28, 2007) to permit all students who meet specified criteria to use AP or IB examinations as substitutes for the State’s HSA in algebra/data analysis and English. The approval permitted some change in the criteria. Beginning that year, Maryland could count as proficient for high school English students who score 3 or higher on the AP English Literature and Composition or the AP English Language examinations or a 5 or higher on the IB English A1 examination. Additionally, Maryland can count as proficient in mathematics students who score 3 or higher on the AP Calculus (AB or BC) or AP Statistics examinations or 5 or higher on the IB Mathematics
Studies SL or IB Mathematics (SL or HL) examinations. ED’s decision letter included this guidance for other states considering adoption of this practice:

The Department recognizes that the AP and IB exams are of high quality and are tied to specific college-level course content. Maryland submitted sufficient evidence that the AP and IB exams meet or exceed the Maryland content standards and the proposed proficiency scores demonstrate that the students taking these substitute assessments have met the corresponding achievement standards. This flexibility agreement will permit Maryland to use the results of the AP and IB assessments in AYP calculations, thereby offering an assessment of college-level course content which may not in every case cover the full breadth of the Maryland high school content standards as required under 34 C.F.R. 200.3(a)(1)(i). This agreement is in effect for four years (the 2006-07 through 2009-10 school years) unless the reauthorization of the ESEA changes the requirements on which it is based. As required by section 9401(b3)(A) of the ESEA, within 30 days of the date of this letter, Maryland must provide all interested school districts with notice of this flexibility agreement and a reasonable opportunity to comment, submit any comments to the Secretary, and provide notice and information to the public about its use of the AP and IB assessments in the manner in which it customarily provides similar notice to the public. Maryland must also annually submit a report under section 9401(e)(1) of the ESEA to the Department at the end of each school year that describes its use of the AP and IB assessments.

Virginia’s approval was also for four years and consistent with the guidance provided to Maryland. However, Virginia received approval to use slightly different achievement scores. Students scoring 2 or higher on the AP English Literature and Composition or 3 or higher on the IB English assessments will be considered proficient on the state’s Reading Literature test. Students who score 2 or higher on the AP Calculus (AB or BC) or 3 or higher on the IB Mathematics Studies, IB Mathematics Methods, or IB Mathematics assessments will be considered proficient on the state’s Algebra I, Geometry, or Algebra II tests.

This year, as noted earlier in this section regarding assessment system changes, Maryland received approval (June 17, 2008) to substitute AP or IB scores for its high school assessment in biology. Students will have to earn a score of 3 or higher on the AP Biology or 5 or higher on the IB Biology SL or Biology HL examinations.

Banking of Test Results
After initial reservations, ED began approving requests in 2004 to “bank” test results used in AYP analyses when students have two or more opportunities to take state tests. As noted in the 2004 paper, “The version of this practice that . . . is acceptable to ED involves the designation of a single test administration as the ‘official’ point at which students are expected to have acquired the knowledge and skills specified in the standards on which the test is based” (Forte Fast & Erpenbach, 2004, p. 4). Thus, if a student takes the test before the official point and achieves a “passing” or proficient score, the score is “banked” and the student does not have to participate in future administrations of the test.

Last year, ED approved requests from Kansas and Oregon to bank scores, and one state, Wyoming, clarified its practice in this area. This year, only one state sought workbook amendments related to the banking of test scores. Wyoming indicated that it would no longer provide for a mid-year assessment window for any of its assessments (approved July 15, 2008). The mid-year assessment was optional but did provide for banking results until year-end whenever students participating in the earlier testing window earned “passing” scores. However, ED’s decision letter seemed to introduce new criteria related to the banking of test scores: “High school English/language arts scores may only be ‘banked’ for accountability
purposes [in Wyoming] if the reading and writing portions are taken in the same testing window and 9th graders will not be permitted to take the English/language arts high school assessments.” State officials subsequently clarified that removal of the option for 9th grade students to take the 11th grade assessment and bank their scores and the restriction on the manner in which English/language arts scores could be banked was based on recommendations from a state task force on assessment. In Wyoming, 10th graders will still be able to take the assessments a year early and have their scores banked for inclusion in AYP decisions that are made at the 11th grade level.

**Recalculation of AYP Using “Best Scores” or Results from Subsequent Test Administrations**

This is another area where ED’s initial decisions precluded the use of “best” scores for students retesting on state assessments. In April 2007, ED removed a regulatory prohibition, originally found at §200.20(c)(3), and permitted the use of results from subsequent test administrations under certain conditions. In 2006, Delaware (June 22) and Virginia (July 27) were granted approval to recalculate AYP using scores for students who retested after the official administration of an assessment. Delaware incorporates recalculated AYP results scores for students who retest after participating in a summer school remediation program. Virginia “include[s] a student's proficient score on the re-test of the high school end-of-course assessments in the calculation of AYP up to Grade 12. The end-of-course assessments include English, Reading, Algebra I, Geometry, and Algebra II in high school.” The combination of ED’s decisions in these two cases appeared to set a precedent. However, in 2007, ED seemed less consistent.

In 2006-07, three states—Arizona, Michigan, and Pennsylvania—sought related amendments. Arizona wanted to use a cohort-group approach for the assessment portions of high school AYP determinations. ED’s June 28, 2007, decision letter to the state, which specified outcomes associated with other amendment requests, did not include any reference to the cohort request. In its submission, Arizona argued that it uses a high-stakes test for graduation. Students may take the test up to four times subsequent to the initial administration in the spring of 10th grade. The state noted that its student identification system permits tracking individual student progress and movement through the high school years.

Michigan resubmitted an amendment request, which ED originally denied in 2003, to recalculate AYP using scores for high school students who retest the statewide assessment to improve their scores. Given the high-stakes nature of this assessment, high school seniors there have been permitted to retest. On July 16, 2007, ED approved Michigan’s use of a student's best score on the new Michigan Merit Exam (MME) for AYP determinations beginning with the graduating class of 2008. This was when the first group of students taking the MME became seniors.

On July 2, 2007, ED denied Pennsylvania’s request to recalculate prior year high school AYP results based on the scores of 12th graders retaking the state’s high school assessments and achieving a higher score. ED’s rationale is instructive in the distinction made between retesting on assessments required for graduation versus those that are not. ED noted:

> Section 1111(b)(3)(C)(ii) of the ESEA requires that a State’s assessments “be aligned with the State's challenging academic content and student academic achievement standards, and provide coherent information about student attainment of such standards.” The PSSA is designed by the State to align to the State’s 11th grade academic content standards. Therefore, even if students score proficient on the 12th grade retest, they have not demonstrated that they meet Pennsylvania’s high school content standards, which are based on what the State deems students should know by 11th grade. In
addition, the ESEA requires the timely notification to parents and the public regarding school and district AYP determinations. Allowing a re-test of 12th graders and re-calculating AYP determinations for the prior year would cause confusion regarding the meaning of those determinations; the statute does not allow for re-correcting of AYP determinations after the fact based upon re-tests in subsequent years.

Three states sought approvals related to the use of “best scores this year.” Georgia, beginning in 2008-09, will count summer re-test scores in its AYP determinations. ED approved the request on June 17, 2008, stating that, “This method may not impinge on Georgia’s ability to make timely AYP determinations.” In Tennessee (approved July 28, 2008), the state will include the use of a “best score” for “students taking multiple test administrations during the same school year.”

Maryland will “use a status model to report assessment results for high school students on the basis of the student’s highest score achieved on the NCLB-required assessments for algebra/data analysis and English regardless of the grade in which the student took the assessment.” The state will phase the change in over two years with scores from grades 9, 10, and 11 counting the first year and from grades 9-12 the second. On June 17, 2008, ED approved the request, stating that:

Maryland will “bank” scores of students taking the high school test. For the 2007-08 school year, Maryland will use the best results on the high school assessment among all 11th-graders when calculating AYP for high schools. Beginning in the 2008-09 school year, Maryland will use the best results on the high school assessment among all 12th-graders.

Miscellaneous Standards and Assessments Amendments
This year, four (vs. five in 2007) states submitted requests related to standards and assessment issues that do not fall into any of the categories above. California reported that its testing contractor “has developed a procedure for adjusting test scores to account for changes in item difficulties resulting from the use of a calculator [by SWDs consistent with IEPs]. This procedure results in separate scoring tables for students who use a calculator and those who do not.” On September 26, 2008, ED approved the use of calculators for SWDs taking the California High School Exit Exam. Further, “calculator use alone would no longer result in students failing to be counted as participants for AYP.” Hawaii reported that it has defined grade-specific and domain-specific performance-level descriptors for standards-based sections of the state’s academic content and performance standards.

Rhode Island, which is bringing on line a new high school level assessment in collaboration with New Hampshire and Vermont, will integrate the writing assessment score at grade 11 into the English language arts score in the same manner previously approved for elementary and middle school grades. Thus, reading will be weighted as 80 percent and writing as 20 percent of the total score (approved June 11, 2008). In Wyoming, the state will make AYP decisions in English/language arts “by taking the weighted average of the number of students scoring proficient or above on the reading test and on the writing test measured against the annual measurable objectives for English/language arts.”

Adequate Yearly Progress (AYP) Model
The requests summarized in this part of Chapter 2 relate to how AYP is calculated and applied. First, amendments related to the application of AYP to schools and districts are described, followed by amendments to AYP indicators. Then, four summaries follow related to growth models, annual
measurable objectives and intermediate goals, states’ strategies for enhancing the reliability of their AYP decisions, and safe harbor calculations.\textsuperscript{24} \textsuperscript{25}

In an interesting development, \textbf{Idaho’s State Board of Education} announced on May 27, 2008 that it would seek a “fresh start” for all of its public schools identified for improvement and facing NCLB sanctions. The action would have had the effect of negating AYP determinations made between 2002 and 2006. The state board argued that Idaho had poorly written academic content standards and used assessments that were not aligned to what teachers were supposed to teach. According to a May 29, 2008, report in the \textit{Idaho Statesman}, “Idaho was fined $103,000 in 2005 after the feds said the state did not have an adequate testing system in place. The board is now using the feds’ own argument against them. If the previous system was poor, the board argues, schools shouldn’t be held accountable under it. ‘How can you say that a school didn’t make progress if your initial measurement was not valid?’” (Roberts, 2008, May 29)? On June 5, ED denied the requesting, commenting:

> Although not framed as such, we construe your request for a “restart” as a waiver under section 9401 of [ESEA] . . . as you are requesting that the Department waive retroactively the statutory requirements applicable in school years 2002-03 through 2005-06 regarding the failure to make AYP and the various remedial measures that a school or district must take for each additional year that it fails to make AYP. . . . A waiver request must, among other things, describe how the waiver will (1) increase the quality of instruction for students; and (2) improve the academic achievement of students. . . . allowing Idaho to restart its timeline would harm students and interfere with these goals by eliminating important actions that would otherwise be taken by schools and districts currently in improvement, corrective action, or restructuring to improve student achievement.

Once again, many states—including \textbf{Alaska, Florida, Missouri, New Hampshire, New York,} and \textbf{Utah}—submitted a number of amendments aimed primarily at clarifying many elements in their accountability workbooks. These were essentially minor in nature and are not summarized in this paper. ED did not appear to deny any of these clarifications. In her November 19, 2007, letter to chief state school officers regarding the submittal of accountability workbook amendment requests for 2007-08, Assistant Secretary Kerri Briggs encouraged states to use this opportunity to include those “required to accurately reflect current practice . . . ”.

\textbf{Adequate Yearly Progress (AYP) Application}

Requests related to AYP application involved modifications to the manner in which AYP is calculated generally; accountability for schools and districts that are new or reorganized; accountability for schools with the highest grade of 1 or 2; modifications to the manner in which AYP decisions are made for small schools; changes in the definition of full academic year; accountability identifications based on the same subject and student group; accountability for schools that use equi-percentile adjustments; and identifications based on scores for only those students served in Title I targeted assistance programs.

\textit{Modifying AYP Calculation Methodologies Generally}

Although the NCLB requirements for calculating AYP are quite prescriptive, states have nevertheless continually sought flexibility from ED with regard to the interpretation of these requirements. And, to some

\textsuperscript{24} Readers will find as interesting additional background reading a paper on educational accountability systems by Perle, Park, & Klau (2007) prepared for the ASR SCASS.

\textsuperscript{25} CCSSO’s ASR SCASS also maintains an annotated list of accountability-related research studies for state leaders at \url{http://www.ccsso.org/content/pdfs/List%20of%20NCLB%20Accountability%20Studies%200108.pdf}
extent, the department has afforded some flexibility (see Appendix C for examples). This year, nine states submitted workbook amendments or clarifications concerning their methodologies for calculating AYP that are best characterized as "general" in nature.

**Arizona** proposed to base AYP decisions on missing a subject area AMO (reading or mathematics), missing the participation requirement in any subject area, or missing the OAI. Missing a subject area target one year and the participation rate target in any subject area the next year would not have moved a school or district forward in terms of improvement requirements. In examining prior years’ data, the state noted that the effect of the proposed modification would have resulted in 25 fewer schools advancing in improvement status last year versus two more that would have moved forward. On July 15, 2008, ED denied the request stating,

This request is contrary to NCLB, which requires a school to meet both the subject AMO and percent participation in order to be considered to have met AYP for a particular subject. See section 1111(b)(1). The purpose of the 95 percent participation requirement is to ensure that schools are held accountable for the performance of all their students, not a select few. This proposal would essentially eliminate such an assurance by allowing a school to alternate between testing all of its students in a subject in an effort to meet the participation requirement and testing a limited number of students in an effort to meet the AMO the next year, thereby creating an incentive to exclude from testing those students that are most likely to have low performance.

**California** ensured that it will make (approved September 26, 2008) school-level AYP determinations for its grade span analyses in the same manner as those used for school district determinations (2-5, 6-8, and 10). The procedure includes:

- The application of safe harbor for both grade spans and numerically significant student subgroups within grade spans.
- The use of a 99% confidence level in the overall AYP determination for grade spans with fewer than 100 valid scores, but not for numerically significant subgroups within grade spans with fewer than 100 valid scores.
- The use of two- or three-year averaging in determining whether or not a grade span or numerically significant subgroup within a grade span has met the participation rate or the percentage proficient or above criteria.

**Hawaii** will base its reading and mathematics proficiency determinations from spring 2007 forward on new cut scores set earlier that year. The revisions were necessary due to the adoption of revised academic content standards and the implementation of new assessments. On September 25, 2008, ED approved Hawaii’s plan “unless the content standards or test design is changed or the Board of Education deems it necessary to revise the cut scores.” In 2007, **Massachusetts** modified its procedures with respect to the collection of race and ethnicity data. For 2008, and beyond, the state clarified its description of racial and ethnic student group reporting, a clarification impacting racial and ethnic student group AYP calculations and reporting (approved August 18, 2008).

**Missouri** submitted numerous minor amendments to clarify its policies and procedures related to AYP calculations (virtually all of which were approved by ED without enumeration). The amendments reflected changes in state policies, changes necessary to address areas of non-compliance reflected in a March
2007 ED monitoring review, and a decision to adopt end-of-course examinations in place of high school assessments beginning in 2008-09. **New Mexico** sought to “apply a t-score rather than a z-score to its confidence interval when calculating AYP.” The state had been applying a 99% confidence interval using a z-score statistic. On July 23, 2008, ED denied the request stating:

> The Department believes that the current New Mexico method is sufficient to minimize the occurrence of a Type II error [identifying a school as meeting a performance target when it did not], which is of the most concern to the Department and New Mexico, and that this proposal is unnecessary and increases the risk of a Type I error [identifying a school as not making AYP when in fact it has], thereby causing schools that may need additional assistance not receiving that assistance.

In **North Carolina**, the state sought approval to hold schools and districts harmless for grades 3-8 reading results due to the administration of a new assessment in this subject area. AYP decisions were to be based on mathematics, participation rate in mathematics, and the OAI for 2007-08 only. The state argued that the staggered release of decisions ED required in 2006 for new mathematics assessments resulted in confusion for both schools and parents. ED denied the proposal on June 6, 2008, stating, “The Department cannot permit a state to base AYP determinations solely on the results of one subject’s assessment. Doing so could lead to students not receiving needed interventions if a school’s reading/language arts results indicate low performance.” In a follow-up visit with the state in summer 2007, ED representatives agreed to “see if there is a better way to work through transitions in test editions when new standards have to be set.”

**Oregon** submitted clarifications regarding how new schools there are constituted and held to NCLB requirements:

> New schools will be held accountable as soon as sufficient data points are available. Two years of data are used to determine if a school has met the annual objective. . . . Students enrolled in newly reconstituted schools will be included in LEA accountability. State policy is that if enrollment of a school changes by more than 40% due to boundary changes, it is considered a new school. Assessment data will be reported on new schools the first year.

**South Dakota** increased the targets that schools and school districts must attain in order to be considered “distinguished.” ED approved the plan stating, “This amendment provides that only student groups meeting the minimum subgroup size of 10 students will be considered in determining whether a school should be recognized as a distinguished school.”

---

26 The Thompson Publishing Group, Inc. keeps track of ED’s Title I monitoring reports. Reports for individual states can be found at [http://www.thompson.com/public/nclb/monreports.html](http://www.thompson.com/public/nclb/monreports.html)
Accountability for New Schools, Reorganized Schools, and Reorganized Districts

States are occasionally confronted with the dilemma of how to apply accountability decisions when new schools are created as a result of restructuring or consolidation. This is a phenomenon that seems to increase yearly as some schools and districts experience progressively more stringent sanctions under NCLB.

Arkansas (approved June 26, 2008) amended its new schools definition “to clarify that new school status may only be applied one time during a three-year period.” Connecticut submitted an amendment intended to clarify the manner in which the state determines if a school is “new” for NCLB purposes (approved July 15, 2008). In Connecticut, a Department of Education committee will consider:

1. Significant changes in the student population served by a school which is not attributed to attrition and/or mobility.
   i. A new entity results in serving students who previously attended other schools in the district.
   ii. Change in student population is 50 percent or more, either an increase or a decrease.
2. A significant reconfiguration of grade spans that would require more than a minor change in grade spans such as a reconfiguration from K-5 to PK-6.
3. Reconfiguration of school attendance areas (redistricting) so that the school serves more than 50 percent different students and/or the school serves significantly different grades.
4. Significant changes in governance structure. To be considered, governance changes must also include significant changes in personnel as well as a strong plan of professional development and support for the changes. For example,
   i. A large school is reconfigured into “smaller learning communities” (SLCs) with each SLC led by a separate administrator.
   ii. The school is managed by an entity, such as a private management company, with a demonstrated record of effectiveness.
   iii. Closing the school and reopening it as a focus or theme school with staff trained and skilled in the focus area.

Missouri will follow a pattern adopted by several other states in defining new schools (approved June 25, 2008). In Missouri,

A school that has been assigned a new building code due to the creation of a new facility, change in grade configuration, etc., will be identified for improvement based upon the AYP status of the prior building or buildings serving the students, unless the school meets the “new school” criteria. A new school is defined as a school with an enrollment change due to grade reconfiguration which results in less than 50% of the previous students or incoming students remaining in the building.

27 In non-regulatory guidance on LEA and school improvement revised and issued on July 21, 2006, ED directed at G-9, pp. 28-29, “If a State has operational rules for determining AYP for new schools that differ from those applied to other schools, the State must amend its accountability plan to provide its definition of a new school and to describe how it determines AYP for new schools including whether and under what circumstances a restructured school can be considered a new school” (ED, July 21, 2006). This statement appears to require an action by a state in a document that has no legal standing; “non-regulatory” guidance typically consists of a set of recommendations that have no regulatory standing. Nevertheless, five states submitted related amendment requests for 2006-07, and another five this year.
New Jersey will employ an “AYP restart” when schools have been restructured (approved July 16, 2008). To meet the restart criteria, a school would have to serve grade levels at least 50% different than the grade levels previously served and the school’s staff could not include 50% or more of the previous staff. In Pennsylvania, a school will be considered new in similar fashion (approved August 18, 2008). In that state, either of two conditions will serve to trigger a “new school” designation:

1. Two or more existing schools consolidate to form a new school and the result is a change in the original population of at least 50 percent; or
2. A school experiences a change in grade configuration that involves at least two tested grades.

West Virginia was also approved (July 15, 2008) for similar criteria to determine when a school becomes a new school. Requests to establish a new school there will have to be submitted “through the county long-range Comprehensive Educational Facilities Plan and approved by the West Virginia Board of Education and the West Virginia School Building Authority.” North Carolina (approved June 6, 2008) is taking a slightly different approach to restarting. In that state, when two or more schools merge, the improvement status will come from the school with the highest average daily membership. When a school is divided, the resulting schools will each take the status of the original school.

In New Mexico, school district achievement results will be applied to newly formed schools whenever a feeder pattern is not available (approved July 23, 2008). That designation will hold for the school “until a tested grade level is established or the students are promoted to a tested grade level where New Mexico’s approved feeder-school methodology is possible.” On June 12, 2008, ED approved North Dakota’s amendment pertaining to reorganized and consolidated school districts. In that state, “The proscribed method joins, compiles, and analyzes multi-year, historical student achievement data to create a newly validated base upon which current and future adequate yearly progress determinations are made for the newly established public school district.” Utah submitted an amendment setting forth a little broader definition of a new school or district. ED approved the request on June 17, 2008, but added to the scope of the state’s definition. According to the Department’s decision letter:

New schools/districts are defined as a school/district that has had no prior enrollment or a school/district with significant (40 percent) change in population or staffing due to boundary changes, school closures, or changes in grade level configurations that now serves a different population such that fair comparisons cannot be made to the prior year’s performance. A new school/district will begin a new base line for purposes of calculating AYP. In subsequent years, data generated by the school/district will be used for making AYP determinations. The goal of 100 percent proficiency by 2013-2014 will be established for new schools/districts.

Accountability for Schools Not Covered by State Assessments and for Small Schools

NCLB requires states (and LEAs) to make AYP decisions for all schools. This poses special challenges when schools do not serve grades covered by a state’s assessment system (e.g., K-2 schools) or where the total enrollment is less than the state’s minimum n for AYP. In regulations issued in December 2002 (regarding educational accountability systems and the calculation of AYP), ED stated that it “will issue non-regulatory guidance to provide examples of methodologies for handling this issue” (p. 71744). That guidance does not appear to have been issued although many states continue to submit related workbook amendments. States generally make AYP decisions for schools that serve only grades not tested (feeder schools typically serving students below grade 3) by using data or results for schools that receive students from the feeder schools. States have also been allowed to apply a variety of strategies for evaluating the AYP of small schools.
Last year, ED approved a total of four amendment requests in this area from three states (New Hampshire, South Carolina, and Tennessee) and denied the request of a fourth (Montana). This year, nine states submitted amendment requests including a re-submittal from Montana.

Arizona clarified (approved July 15, 2008) that it intends to make AYP determinations for K-2 schools “based on the AYP determination of the third grade of a designated school to which a plurality of the students of the K-2 school matriculate.” According to the state, “The essential impact of this proposed change is to remove from the evaluation students who move outside the district of the K-2 school.”

California (approved September 26, 2008) will “pair and share” elementary schools such that K-1 schools “will be assigned values derived solely from the second grade test results of schools with which they are paired.” Currently, the state assigns aggregated results of all tested grades from the receiving to sending school. California argued that “Test results from the second grade only would be a more immediate and dependable indicator of performance.” In Colorado, the state will calculate AYP for K-2 schools as follows (approved August 4, 2008):

1. K-1 and K-2 school AYP will be determined using the third grade reading and math scores of students previously enrolled at the school.
2. K-1 and K-2 schools will be held to elementary school AYP targets for accountability purposes.
3. All schools will be expected to yield annual results that meet the requirement of 100 percent proficiency in reading and math by 2013-14.

Idaho has 51 schools with fewer than the state’s minimum n in tested grades. The state had previously made AYP decisions by averaging test results over a three-year period. On June 30, 2008, ED approved Idaho’s request to use the highest of the current year’s score or the three-year average when making AYP determinations. Missouri will link K-2 schools “with and receive AYP determinations on the basis of test results of the schools their students attend in subsequent years.” ED also approved (July 23, 2008) New Mexico’s criteria for the determination of when a reorganized school in fact becomes a new school. According to the approval:

A “new” school will receive the school improvement designation of a sending school that makes up 61 percent of the new school or, in the case of a reorganized school that is populated by students from various sending schools, if no one makes up 61 percent of the reorganized school, but which, collectively, account for 61 percent or more of the students enrolled at the reorganized school that school will receive the AYP designation that occurs most frequently of all the sending schools.

In North Carolina, the state is planning to change its processes for making AYP determinations for schools with two or fewer students meeting FAY requirements. North Carolina has used site visits to make these determinations in the past. On June 6, 2008, ED approved the state’s plan “to permit the LEA to provide a justification for the AYP designation” in place of the site visit by a team of educators. Pennsylvania will, for AYP determinations, use district-wide results of the grade into which students enter when they “feed” from schools not administering state assessments because of school grade configurations (approved August 18, 2008). ED’s decision letter also affirmed:

However, because charter schools do not necessarily feed into any particular district, charter schools that do not administer statewide assessments because of their grade configurations will determine their AYP status based on their other academic indicator: graduation rate for those charter schools that have a graduating class or attendance rate for those charter schools that do not have a graduating class.
Wyoming (approved July 15, 2008) will pair any newly opened K-2 school with an existing elementary school that includes grades tested for AYP purposes.

Montana again sought to expand its “small school review” (SSR) process for schools serving fewer than 30 students to schools that serve 30 or more students. In making its latest request, Montana argued:

The process currently is used to make accountability determinations for districts and schools that otherwise would not be included in the accountability system under the No Child Left Behind Act (NCLB) because of the small numbers of students enrolled. We believe strongly that use of that process, which examines other confirmatory data in addition to achievement test scores and attendance or graduation rates, results in the most accurate determinations regarding the performance of Montana’s schools and districts, and allows educational interventions (i.e., rewards and consequences) to be targeted appropriately. Simply put, that process makes sense for Montana, and its extension to all districts and schools will result in a uniform accountability system that will be understood and supported by our citizens.

ED denied (August 2, 2007) the state’s initial request and also, on July 16, 2008, the re-submittal of the request. The Department’s 2008 decision statement was almost identical to that provided Montana last year:

This request is inconsistent with section 1111(b)(2)(I) of the ESEA, which requires a state to measure AYP for all students and student subgroups according to a state's annual measurable objectives, which must be defined consistent with statutory provisions that govern setting starting points and intermediate goals over a 12-year timeline. This requirement is tempered only for groups that fall below the state's minimum group size needed to yield statistically reliable information. Montana's SSR, although approved for use with small schools, bases AYP determinations on a subjective scoring rubric and does not emphasize the performance of subgroups. In addition, we note that the vast majority of students in the state (75 percent) are enrolled in eight school districts, with the largest, Billings, enrolling 15,000 students. It is not necessary or appropriate for Montana to use its SSR for large districts and schools to ensure valid and reliable results.

Both last year and this year, Montana, perhaps anticipating ED’s denial, included another request to be considered, involving increasing the minimum group size to 40 students. ED denied that request as well last year and again this year.

Identifying School Districts for Improvement

In 2004, ED approved a new approach as to how school districts could be identified for improvement (see Forte Fast & Erpenbach, 2004, p. 9). Initially referred to as the “Tennessee model” because the state was the first to advance the concept, districts would be identified for improvement only if they miss an AYP target for two consecutive years in the same subject across all grade levels (elementary, middle, and high school). The following year, ED clarified that this included missing the OAI at all three grade levels. While at least 18 states sought related amendments in 2005, only seven did in 2006, and none in 2007.

This year, three states submitted an accountability workbook amendment concerning how school districts are identified for improvement. In Rhode Island (approved June 11, 2008), the state adopted Tennessee’s model. In Rhode Island:

Districts are not “identified for improvement” until they have missed any math target for two consecutive years, any English language arts target for two consecutive years, or the attendance rate
or graduation rate target for two consecutive years. Missing a participation rate target is equivalent to missing the corresponding content target area.

The identification of school districts for improvement can also be impacted in cases where students are served in alternate settings or settings away from their “home” schools or districts. Oklahoma requested an amendment this year specific to such a case (approved June 17, 2008). In this instance, the state clarified that “students who are placed in a school site or district outside their enrollment area (such as incarcerated youth and students placed by court order), and thus identified as ‘other placement,’” will be included “in the state’s virtual statewide district for adequate yearly progress . . . and graduation rate calculations.” South Dakota clarified (approved July 15, 2008) the responsibility of school districts related to students whose instruction is provided “out-of-district.” According to ED’s decision letter, “when a student has been assigned out of district and is enrolled in a South Dakota school operated to serve the special needs of a student (e.g., special education or alternative programs), the student will be counted in his or her district of residence for accountability purposes.”

**Full Academic Year Definition**

States customarily calculate percent proficient based on the achievement of only those students who have been enrolled in a school or district for a full academic year (FAY). Under NCLB, states define FAY and ED has given states considerable leeway in their definitions provided that these do not go beyond one calendar year. In recent years, the number of states submitting amendments to or clarifications of their FAY definitions has varied—eight in 2006, one in 2007, and nine this year.

Georgia re-submitted a proposal that was denied by ED last year: that the SWDs student group will consist of students reported to be receiving special education services from the fall enrollment count (first Tuesday in October) through the spring testing window. This ensures that the group reflects the achievement of “children served by special education programs and services during the school year rather than students who might have only received services for a matter of days.” ED’s position in its June 14, 2007, letter was that the change “could create a negative incentive to identify students as SWDs late in order not to be counted in the subgroup.” Further, “The length of time a student receives related services has no impact on whether a student meets the definition.” The department again denied the proposal on June 17, 2008, for virtually identical reasons.

Hawaii clarified (approved September 25, 2008) that it will define FAY “to mean continuous enrollment from March 1st of one school year to March 1st of the next school year . . . ”. However, in approving the clarification for only 2007-08 purposes (September 25, 2008), ED added a stipulation not seen in previous approvals of similar definitions for other states:

> Please note, however, that the problem in Hawaii where students might meet the “full academic year” definition in one school and be tested in another school must be resolved for the 2008-09 school year. Any student who meets the full academic year definition in a school should be assessed in that school. The student has met the state’s definition of receiving a full year’s instruction and thus should be included in the school’s accountability determination. Hawaii must address discrepancies between its testing window and full academic year definition.

Missouri will change from enrolled on the day that initial enrollment (last Wednesday in September annually) is reported to the state to enrollment on that day through the statewide assessment window (approved June 25, 2008). Louisiana proposed to add “more specificity to subgroup labels as they are used in accountability for establishing FAY. Students who are newly labeled as members of a specific subgroup between October 1 and the assessment period shall be considered members of that subgroup
for the entire academic year.” ED’s September 24, 2008, decision letter to the state did not mention this proposal although it did note that “Louisiana clarified language throughout its workbook to reflect the state’s current practice and policy.”

Under Nebraska’s current assessment system, “assessments are taken in the classroom throughout the school year.” The state clarified that the definition of FAY is “enrolled from the last Friday in September until the end of the school year or the end of the assessments,” whichever comes first. In Oklahoma, the state amended its FAY definition to be consistent at all levels, grade 3 through high school. Oklahoma will now use the high school definition at all levels (approved June 17, 2008)—“A student has been continuously enrolled beginning with the first ten days of the school year and has not experienced an enrollment lapse of ten or more consecutive days.”

South Dakota will change its FAY definition for SWDs from October 1 through the statewide testing window to December 1 through the statewide testing window. December 1 is the state’s Child Count Day and the change will be more inclusive for SWDs. The FAY definition will remain the same for all student groups (approved July 15, 2008). Utah sought to change its FAY definition for the state and school districts to “on or before May of the previous school year and continuously enrolled until the testing period of the current year.” Currently, the state applies the definition used for schools—membership for not less than 160 days. ED did not approve the amendment according to a state official without mention of the matter in the state’s June 17, 2008, decision letter. However, the official clarified that the denial stemmed from what would have been the use of two FAY definitions—one for schools and another for districts and the state. Wyoming (approved July 15, 2008) will now define FAY as enrollment from October 1 through the 15th day of the annual testing window. The state will no longer provide a mid-year testing window.

Identification for Improvement Based on Same Student Group and Same Subject
As reported in each of the prior papers in this series (Erpenbach, Forte Fast, & Potts, 2003, Forte Fast & Erpenbach, 2004, Erpenbach & Forte, 2005; Forte & Erpenbach, 2006, and Erpenbach & Forte, 2007), ED has consistently rejected proposals to limit identification for improvement to cases where the same student group misses the same AYP indicator in each of two consecutive years (misses in the same “cell” two years in a row) even though the law and regulations are silent on this matter. More recently, Secretary Spellings announced ED’s proposed new regulations in nine areas of NCLB, including the prohibition of identification for improvement based on the same student group and subject (Spellings, 2008, April 23).

In 2007, five states—Michigan, Oklahoma, Pennsylvania, Utah, and Washington—sought approval to base AYP decisions on the “same cell” model. ED denied all five requests. This year, four of those states re-submitted their requests (Oklahoma, Pennsylvania, Utah, and Washington).

It is worth noting that states’ requests and rationales in this area have been relatively consistent with the following excerpt from the 2005 paper (Erpenbach & Forte, 2005, p. 17).

After working with these schools [those identified for improvement based on missing AYP for any subgroup in the same subject for two consecutive years] for a number of years, [our] experience has shown that it is difficult for schools to develop meaningful school improvement plans when they face a moving target. Schools identified for improvement due to a different subgroup each year have a difficult time determining whether if the problem is due to an anomaly, is a function of the definition, or whether it is truly a problem that can be traced to academic achievement and their program. A moving target, such as one in schools identified by a different subgroup each year, leads to a sporadic and inconsistent
implementation of school improvement efforts. Applying sanctions only to schools consistently failing to make AYP for the same subgroup and subject will ensure that school improvement efforts are focused on schools with systemic issues of academic underperformance. This lends a great deal of credibility to the formal school improvement identification process and removes the convenient excuse that the identification was a function of the definition.

Similarly, Virginia proposed in 2006 that identification for improvement be based on same student groups missing the same AYP target in two consecutive years, arguing that,

USED [U.S. Department of Education] regulations permit states to identify for improvement only those schools that fail to meet AYP for two consecutive years in the same subject, but prohibit states from treating subgroups the same way. This model raises reliability concerns given that many groups . . . could fail to demonstrate AYP for any given year. This policy also fails to recognize the different educational challenges and interventions that may be appropriate in cases where different subgroups fail to make AYP. Identifying schools for improvement based on not making AYP for two consecutive years in the same subject and same subgroup will target resources to the particular subgroup(s) that need them the most.

Prior to 2007, the only rationale ED provided in writing for its denials of these requests was included in a letter to West Virginia (December 31, 2003): “This identification [same cell two consecutive years] procedure is not consistent with the statute or regulations.” No citation or support was provided for this statement. However, in denying similar requests from four states last year, ED did include a written rationale in each case. The response to Oklahoma is illustrative.

In rejecting Oklahoma’s request last year (June 8, 2007) and again this year (June 17, 2008), ED essentially noted both times that it has previously pointed out to states the flexibility found in section 1111(b)(2) that permits a state to determine schools in need of improvement on the basis of not making AYP in the same subject for two consecutive years. This, ED asserts, is because other provisions in the law treat reading and mathematics independently. On the other hand, ED notes that the law does not treat student groups28 independently; thus, ED concludes that the law does not afford schools and districts similar flexibility to identify them for improvement on the basis of same student group performance for two consecutive years. Washington, in re-submitting its amendment proposal this year, also found it denied once again by ED (August 18, 2008). The department’s rationale was essentially the same as that given to Oklahoma.

Clearly, many states continue to believe that just the opposite conclusion could be drawn—it is the performance of a given student group in a given subject over two consecutive years that should be of the utmost concern. As previously noted in this and earlier papers in this series, the statutes and regulations are silent on this issue.

In its 2008 re-submittal, Pennsylvania again argued to not advance schools and districts in the improvement cycle “unless they fail to meet AYP for the same subject and same subgroup for two consecutive years.” The state further argued that, “The failure of one subgroup to improve in a given subject indicates a school’s/district’s need to more fully identify approaches to improve performance; thus, the school/district should proceed in AYP status. However, the failure of different subgroups from one

---

28 The term, “student group,” is used more commonly now in place of the term “subgroups.”
year to the next is too often based on volatility of the subgroup’s population.” If this practice had been in place in 2007, Pennsylvania estimates that of 298 schools that proceeded to the next level of improvement, only 135 (45.3%) would have moved forward in terms of sanctions. Nevertheless, ED again turned down the amendment (August 18, 2008) for the same reasons given last year as well as those given to Oklahoma and Washington this year.

Utah added sampling or other sources of errors to its rationale in supporting a same-cell for two consecutive year’s proposal:

> If a school has a student group that fails to meet its AMO in the same content area for two consecutive years, we can be confident that the school has a shortcoming that must be addressed. While it is a concern when any student group fails to meet its AMO in a particular content area in a given year, we recognize that this yearly failure could be due to sampling or other sources of error, therefore we want to be confident of the need before implementing specific sanctions. Identifying schools after the same subgroup fails to make AYP in the same subject for two years provides a much clearer illustration of where the school needs attention. The school can focus on examining the special needs of the students, improve their instruction, and effect change. This approach is consistent with the NCLB goal of successfully remediating deficiencies to close the achievement gap, and will reduce the potential for falsely concluding that a school is not meeting AYP requirements.

On June 17, 2008, ED again denied Utah’s proposed amendment. The department’s decision was similar to its response to Oklahoma, concluding that:

> The intent of school identification is not to lay blame on a particular group of students, as the “same subgroup/same subject” proposal would do, but to identify the instructional and academic areas that need to be improved. A school that is identified for improvement should look to specific instructional remedies in the subject area, other indicator, or participation rate that resulted in its identification.

**Use of Equi-percentile Adjustments** for AYP Determinations

Georgia was approved (June 17, 2008) for a second year to continue a previous approval related to AYP determinations as the state transitions from one set of academic content standards and assessments to revised standards and aligned assessments. As described in its 2007 submittal, to make AYP determinations Georgia “will equate the QCC (older) to GPS (newer) assessment results in grades and subjects where appropriate, using an equi-percentile adjustment for multi-year averaging, safe harbor, and second indicator calculations. This method will be applied for two years across each subject and grade.” The method includes safe harbor reviews and may be applied for two years per subject per grade.

In April 2008, ED approved New Hampshire’s continued use of the equi-percentile method for determining safe harbor as the state transitions to new assessments in grades 3-8 and 11.

---

29 According to Scott Marion at the Center for Assessment (personal communication, July 29, 2006), “When we change tests in ways that prevent us from doing a formal equating, our best bet is to simply project the growth trend from the past test on the new test scale (e.g., if the scores have been increasing 2% per year, we would project a score [mean score or % proficient] 2% higher on the new test). This is only used to allow us to do things like calculate safe harbor or roughly estimate trends. If the tests changed enough not to conduct formal equating, we will re-set the achievement standards and these will not be affected (except as impact data) by the equi-percentile linking.”
Exempting Schools and Districts Not Receiving Title I Funds
Under NCLB, states can choose to apply the law’s sanctions to all public schools and districts or only to those receiving Title I funds. From time to time, a few states have put forth amendments around this provision that have been approved by ED. This year, two states submitted such amendments.

In **California**, the state clarified that (approved September 26, 2008):

An LEA or school that is identified for PI [program improvement] but ceases to receive Title I funds is not required to continue PI activities. However, it retains its former PI status up to three years, unless it makes AYP two consecutive times during the period without funds. If the LEA or school once again receives funds within the three-year period and has not made AYP two consecutive times during the period without funds, it will be required to implement the same level of PI activities as when it stopped receiving Title I funds.

**Missouri** also clarified that, “Non-Title I Schools and Districts must write a school improvement plan to address areas of deficiency. Non-Title I districts and schools may also lose eligibility for a waiver of their on-site Missouri School Improvement Program (MSIP) review.” ED approved the clarification on June 25, 2008.

**Adequate Yearly Progress (AYP) Indicators**
AYP indicators are the statistics states use in their AYP models and include percent proficient for reading or language arts and mathematics; participation rates for reading or language arts and mathematics assessments; and at least one other academic indicator of student performance, which must include graduation rate at the high school level.

**Percent Proficient—Use of Indexing**
In 2006, 10 states requested permission to use an index in lieu of percent proficient. Three states that submitted related amendment requests in 2006 (New Hampshire, Washington, and Wyoming) re-submitted them last year—one had been denied in conversations between state and ED representatives; one had been approved; and ED verbally approved the other according to state officials. This year, only two states submitted an educational accountability workbook amendment pertaining to the use of indexing for AYP determinations. Also, ED agreed to extend New Hampshire’s conditional approval first granted in 2006.

Typically, indexing is based on one of two weighting schemes: either weighting achievement levels or weighting by student enrollment across grades. Until last year, ED generally permitted states to use either type of indexing as long as separate measures were calculated for reading or language arts and mathematics. The achievement-weighted type of index has carried the additional restriction that higher performance (e.g., proportion of students scoring in the advanced achievement level) cannot compensate for lower performance (e.g., proportion of students scoring in the below basic achievement level). In practice, this has meant that the highest weight applies to all scores at or above proficient; states could not weight achievement in an advanced (above proficient) level higher than achievement in the proficient level. Given the pattern of responses and non-responses in 2006, it seemed that ED was backing away
from further approval of changes related to indexing (Erpenbach & Forte, 2007, p. 25). That opinion was borne out in ED’s response to New Hampshire’s 2007 amendment proposal. 30

ED approved (May 23, 2007), for one year only, New Hampshire’s request to use the index system orally approved in 2005-06 for AYP determinations in the 2006-07 school year. Under this model, the state “allocate[s] 100 index points for performance at the Proficient with Distinction and Proficient level, 80 points for the upper portion of Partially Proficient, 60 points for the lower portion of Partially Proficient, 40 points for the upper portion of Substantially Below Proficient, 20 points for the lower portion of Substantially Below Proficient, and 0 points for No Score.” In its decision letter, ED commented, “New Hampshire may only use the index for the AYP determinations based on assessments administered in the 2006-07 school year. New Hampshire must submit a new index, based upon additional guidance from the Department to be released in summer 2007 [yet to be released], for approval by the Department before making AYP determinations for assessments administered in the 2007-08 school year.”31 On April 25, 2008, ED approved the state’s continued use of the same indexing model.

Washington, whose 2006 indexing proposal was verbally denied by ED, submitted another proposal last year that was approved in writing (June 4, 2007). The model, likely more appropriately characterized as multi-grade averaging, is similar to that approved in 2006 for Iowa (and earlier for Alabama and Mississippi) and involves the use of a proficiency index for a school with multiple grade spans:

A school with multiple grade-spans will receive a single AYP determination using the proficiency index. The index will be calculated through several steps: 1) determining the difference between percent proficient and the annual measurable objective (AMO) for each grade; 2) establishing a proficiency index weighting constant by dividing the number of students in a grade by the total number of students within the school; 3) calculating the difference between the percent proficient and the AMO multiplied by the proficiency index weighting constant. The proficiency index for the school is the sum of all individual grade-level proficiency index components. A proficiency index of zero or higher indicates that the AMO has been met by a subgroup in the school.

This year, New Jersey reported that it is exploring “methods that will improve validity and reliability of the state’s accountability system.” This would include the use of an index model. The state did not provide details but did indicate that those would be included in their final 2008 accountability workbook submission. In New Mexico, the state has decided to forego its earlier approval to use performance

30 Proposed new regulations announced by Secretary Spellings on April 23, 2008, would require states to explain in their accountability workbooks how minimum student group sizes and other AYP definition components such as indexes combine to provide statistically reliable information.
31 On June 27, 2007, Assistant ED Secretary Kerri Briggs sent a letter to chief state school officers in states approved to use proficiency indexing in AYP determinations inviting them to send representatives to a meeting (held September 7, 2007) to analyze “the impact of performances indexes on AYP determinations to ensure that they support accountability.” Briggs further stated that, “The Department will use the guidance from this meeting to inform any necessary changes to the current use of performance indexes by States, as well as any future proposals from States that wish to use an index in determining AYP.” As a result of the meeting, and based on comments received, guidance on the use of performance indexes in AYP decisions was drafted by an external consultant at the request of ED. However, in a November 16, 2007, email from an ED official to states participating in the September meeting, it was noted that the guidance remains under review and is unlikely to be released while the department re-thinks the matter (Rooney, 2007, November 16).
indexing because, in practice, it has not applied indexing when making AYP determinations (approved July 23, 2008).

**Participation Rate**

The most common amendment request regarding participation rates in 2005-06 involved a clarification to count as non-participants students with invalid test scores; in other words, states cannot count as a participant a student who attempted a test but whose score was invalidated. Other amendments involved averaging participation rates over two or three years whenever the current year rate did not meet the 95% participation requirement. Last year, only three states submitted requests for accountability workbook amendments related to the determination of participation rates (Mississippi, Oregon, and Wyoming).

This year, seven states submitted a participation rate amendment or clarification. Florida will use two-year averaging whenever the current year rate is less than 95% (approved August 18, 2008). Hawaii clarified (approved September 25, 2008) that students exempted from state assessments as a result of written parent requests will be counted as non-participants for participation rate determinations. ED’s approval included two caveats:

1. No school may solicit or encourage a written exemption request on behalf of any child or group of children; and
2. Every school is expected to inform parents upon parent request of the procedures for submitting a written exemption request.

Nebraska completed a new data system, the Nebraska Student Staff Record System. As a result, the state participation rate:

Will be calculated so that the count of not assessed students will be taken from the total of all students enrolled on the last day of the school year and reported on the “Student Snapshot” . . . as having:

- Parent waivers;
- Unapproved modifications or out-of-level assessments; or
- Designations of “N” (Not Assessed) for all achievement levels on every standard. A student’s achievement level is a required field for each standard report for each student. The possible achievement level designations are: 1 – Beginning; 2 – Progressing; 3 – Proficient; 4 – Advanced; M – Moved; and N – Not assessed.

New Jersey will use multi-year averaging—up to three years—whenever a student group’s rate drops below 95% (approved July 16, 2008). Oklahoma clarified (approved June 17, 2008) that the state “will calculate participation rates on its reading/language arts and mathematics assessments separately by subject.” Beginning in 2008-09, Wyoming “will calculate participation rates using enrollment data collected 15 days into the testing window” (approved July 15, 2008).

Colorado (approved August 4, 2008), provided extensive clarification regarding procedures used to determine whether students have been validly assessed and should be included in participation rate calculations. In that state,

- Students who were coded as “test deferred due to language” . . . are counted as non-participants, unless they are recently arrived English language learners who have an overall Colorado English Language Assessment (CELA) score—such students are included as participants for the reading
assessment. Additionally, students who have been in the U.S. for more than three years are counted as non-participants if they take the Lectura (Spanish) assessment. Colorado’s definition of AYP requires that 95 percent of all students in the school or school district as a whole, participate in general (CSAP) or alternate (CSAPA) state assessments.

- If students are coded with a test invalidation code of 4 (parent refusal), 5 (test not completed), 7 (extreme frustration), 8 (non-approved accommodation or modification), 9 (misadministration), or B (District Ed. Services), they are counted as non-participants. They are excluded from the performance and other indicator calculations.

- If a student is coded as test invalidation B (District Ed. Services) and has suffered a significant medical emergency that prevents him or her from attending school and participating in the assessment during the entire testing window, including the make-up dates, the district may appeal the record and have it excluded from participation calculations. Determination that such students have been determined by a medical practitioner to be incapacitated to the extent they are unable to participate in the appropriate State assessment must be included with the appeal.

Other Academic Indicators

Elementary/Middle School Levels

Since 2002-03, few states have requested changes related to their other academic indicators (OAIs) at the elementary and middle school levels, but four requested such a change last year with mixed success; two were approved (New York and Virginia) and two were denied (California and Oregon).

This year, three states and the District of Columbia proposed amendments related to OAIs at the elementary and middle school levels. The District of Columbia received approval (July 28, 2008) for its decision to revise the target for attendance rate calculations. The daily attendance target will be 90% with schools or districts below that required to make at least a 1% increase from one year to the next. In Iowa, the state will use attendance rate (with three-year averaging) at the elementary/middle school level for school as well as school district AYP measures. Iowa had used the indicator at only the district level prior to 2008.

New York will use science as the OAI for elementary and middle schools (approved August 18, 2008). Oregon submitted clarifications regarding how it will calculate attendance rate—the elementary and middle school OAI. The state wanted to continue to use its standard of at least 92% attendance over two years combined or the current year only. However, the state wanted to make the calculation based on:

The attendance rate of students who attended more than 50% of the session days at the institution they were attending on the first school day in May, excluding students who are in a district special education program attending less than 40% of their time in a general education classroom. The district level attendance rate would be calculated similar to the school rate but would include students in district special education programs.

ED denied the request on May 19, 2008. The department did not provide a specific reason. Instead it stated only that the “amendment does not comply with the statute or regulations” and cited section 1111(b)(C)(vi) of Title I.

Graduation Rate—High School Level

A debate started in 2006 has continued ever since regarding the accuracy of states’ reported graduation rates; especially how those rates are determined. The National Governors Association (NGA) enlisted all 50 governors to commit to ensure that the manner in which their states determine and report graduation
rates will be based on the number of entering 9th graders and those who graduate four years later. More recently, Secretary Spellings announced that ED would seek new regulations that would establish a uniform definition of graduation rate, establish a statewide graduation goal, define how schools and districts “demonstrate continuous and substantial improvement from year to year,” and require disaggregated rates as a part of AYP calculations (Spellings, 2008, April 23).

This year, 15 states and the District of Columbia submitted graduation rate amendment requests and another clarified a change it had requested in 2007. This compares to 14 states in 2007, 12 in 2006, 11 in 2005, and five in 2004. This year’s requests included changes to graduation rate targets, how graduation rates are calculated, recognizing special education diplomas in graduation rates, use of interim graduation rates, use of a three-year cohort instead of four years, and inclusion of students graduating in more than four years.

Alaska will establish “a system of reviews to address . . . small school cohort issues.” The state noted, “Often a single student dropping out of school will cause a school to miss adequate yearly progress.” The amendment will provide three levels of review—averaging rates over three years; cohort aggregation over time to meet minimum group size; and, review and evaluation of trends whenever averaging or aggregation results in insufficient data. On February 12, 2008, ED approved the “aggregation of the graduation cohort over two consecutive years until the cohort reaches the state’s minimum group size.” With respect to the absence of sufficient data, ED approved Alaska’s evaluation of “the trend in graduation rates over the past four consecutive years, including the current year. If the graduation rate has increased over the past four years, the school will be counted as meeting the graduation rate by making progress towards the graduation rate target.” Colorado, beginning next year, will use a longitudinal cohort graduation rate similar to that approved by NGA in AYP decisions (approved August 4, 2008).

The District of Columbia received approval (July 28, 2008) for its decision to revise the targets for graduation rate calculations. The district will use a target of 69.9% for graduation with schools or districts below that required to make at least a 1% increase from one year to the next.

Georgia submitted two amendment requests pertaining to graduation rate determinations. Under the first, a re-submittal of a request ED denied last year (June 14, 2007), the state again sought to have its special education diploma included as a regular education diploma for AYP purposes. Georgia argued that some SWDs “are unable to pursue a regular diploma as stipulated in their IEP . . . and have completed 13 years or more of school . . . . Their successes should be included in Georgia’s graduation rate calculations.” ED denied the request on June 17, 2008, citing the requirements of section 1111(b)(2)(C)(vi) that graduation rate must be based on receipt of a “regular diploma” aligned to grade-level content standards.

Under Georgia’s second graduation rate amendment, the state will “use course completion rates and end-of-course test results in place of graduation rate . . . for alternative schools with high school grades.” The state argued that students served in these schools are “off track and need additional time to earn a regular diploma. We want the flexibility to judge these students and schools on their individual semester by semester success” (approved June 17, 2008). Georgia has 21 alternative high schools not offering a

32 In a 2008 report (Curran, Hartney, & Pika), NGA concludes that states are “on track” to publicly report their graduation rates using a common formula by 2012. However, according to the report, only 16 states are using the NGA compact formula with five more planning to report graduation rates according to this formula at the end of 2008, eight more in 2009, and nine more in 2010.

Statewide Educational Accountability Systems Under the NCLB Act
A Report on 2008 Amendments to State Plans 45
high school diploma. The state will “continue to use graduation rate as the second indicator in alternative high schools that do offer diplomas.”

**Iowa** clarified that it will include graduation rates at both the school and district level when making AYP determinations. For 2007-08, the state will use a three-year cohort rate (modified NGA model) and then move to a four-year cohort rate beginning in 2008-09. On July 28, 2008, ED approved Iowa’s revised graduation rate targets but made no mention of the plan to initially use a three-year cohort rate. Because of recently passed state legislation, **Maryland** initially planned in 2007 to implement an interim on-time graduation rate while preparing to put into place a longitudinal data system with a unique identifier for all students. The interim rate was to be used for 2006-07 AYP decisions with the permanent on-time rate coming online in 2010-11. Prior to 2006-07, the State used a “leaver rate” for AYP. However, in August 2007, Maryland withdrew the request due to technical issues in attempting to apply “the interim on-time rate to an accountability purpose . . . [and] the complexity of constructing such files without a unique student identifier.”

**Massachusetts** raised its graduation rate from 55 to 60% using a minimum four-year standard. The state also proposed that a student group “not meeting the four-year standard may make AYP by meeting a 65 percent standard applied to the group’s five-year cohort . . . rate or by increasing by two percentage points the group’s four-year . . . rate . . . compared to the four-year . . . rate for the previous year.” ED rejected the state’s request to include a standard applied to the five-year cohort rates, but, on August 18, 2008, approved the other two components of the plan for the 2007-08 school year only, stating, “To obtain approval for subsequent years, the Department expects that Massachusetts provide a more challenging graduation rate target in future years.”

**Michigan** sought approval to permit more than four years to be used in defining the “standard number of years” for graduation,33 “on a case by case basis, for alternative education programs and schools.” The state’s argument follows:

In the 2007 amendment, Michigan identified the following groups of students where the “standard number of years” for graduation may be five years, based on a case-by-case review of specific information about individual students. The specific groups of students identified in the 2007 amendment were:

- Students with Disabilities;
- English Language Learners;
- Students who have been unable to attend school full time or for a full year due to health problems or mandatory expulsion; and
- Students in Migratory Status.

The request to include students in migratory status was denied, but the other three categories were accepted on a case-by-case basis.

---

33 Michigan, which last year adopted NGA’s cohort model for calculating graduation recently reported that the “proportion of students graduating went from 85.8 percent in 2005-06 to 75.5 percent in 2006-07” (Aarons, 2008, p. 4).
The requirements of Michigan’s Merit Curriculum are rigorous and the state statute includes a provision to create a personal curriculum modification that could result in a student taking extra time to meet the credit requirements.

Data show that, while there are many schools that have students in alternative programs, the students in those programs account for less than 4% of the high school student enrollment.

<table>
<thead>
<tr>
<th></th>
<th>Number of Schools</th>
<th>Percent of Schools</th>
<th>Number of Students</th>
<th>Percent of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>All high schools</td>
<td>1,128</td>
<td>100%</td>
<td>538,642</td>
<td>100%</td>
</tr>
<tr>
<td>High schools with</td>
<td>366</td>
<td>32%</td>
<td>18,263</td>
<td>3.4%</td>
</tr>
<tr>
<td>Alternative students</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data from Center for Educational Performance and Information

The four-year on-time graduation rate does not accommodate students who leave school and seek to return. It is important to not exclude any student who is graduating; this is especially important since graduation rate is a critical factor in determining AYP. By including this additional group in an extended graduation cohort, schools and districts will benefit from the good work they are doing in helping students stay in school to complete their education. Michigan seeks to reduce barriers for schools serving such students who remain in school beyond 4 years. Michigan proposes to allow case-by-case consideration of extending the “standard number of years” for graduation for individual students that are attending alternative education schools, are behind in credit-earning toward graduation, but will graduate earning all required Michigan Merit Curriculum credits. Some programs that serve alternative education students may be labeled as “dropout recovery.” For alternative education programs that serve as dropout prevention or dropout recovery, we will include extra time, through the year a student turns 20, to allow students to complete all required credits.

ED denied the amendment on July 15, 2008, stating, “This request is not aligned with the statute or regulations.” Further, “The Department understands that some students with disabilities or recently arrived limited English proficient students who attend a school in the United States for the first time may need some additional time to attain a diploma and has accordingly granted Michigan flexibility in determining graduation rates for these students.” Therefore, “Since Michigan’s request is not aligned with the current regulations, the Department cannot approve this amendment.”

As a result of a Title I monitoring visit in 2007, ED directed (September 3, 2008) that Nebraska would have to address issues related to the use of graduation rates in AYP decisions there. ED specifically cited “taking into account progress on the other academic indicator (writing and graduation rate) for subgroups within a school or district when applying Safe Harbor calculations to determine AYP, and making AYP decisions for all schools.” New Hampshire will add a 75% graduation rate or an improvement over the previous school year to its AYP requirements (approved, April 25, 2008). New Mexico received approval (July 23, 2008) to “delay the use of a cohort graduation rate for one additional year in order to capture summer graduates in 2008.” In that state, “The rate will be reported in 2008-09 in order to include summer 2008 graduates, developing a one-year lag.” North Carolina wanted to use five years as the “standard number of years” to calculate cohort graduation rates as an AYP measure. The state indicated that it would “continue calculating and reporting a 4-year rate but would use the 5-year rate as the measure for the other academic indicator and for determining if safe harbor can be used at the high school level.” On June 6, 2008, ED denied the request stating, “The Department does not believe that five
years is the standard number of years for the vast majority of students to obtain a regular high school diploma.”

Rhode Island indicated that it will move to an on-time, four-year graduation rate “computed by tracking the same students from year-to-year through high school using their student identification numbers.” This will involve re-computing a graduation rate baseline employing the new method. The state subsequently clarified for ED that since it now has a fall assessment program for high schools, the graduation rate from the end of the prior school year will be used for AYP calculations (approved June 11, 2008). South Dakota will make a similar change this year. According to the state, “Graduation rates for all high schools will be calculated in July, the mean rate determined, and the target set at one standard deviation below the mean.” ED’s July 15, 2008, decision letter stated only that, “South Dakota will implement a graduation rate that is consistent with the compact created by the National Governors Association.”

Tennessee proposed that its graduation rate determination would consist of meeting at least one of four options:

1. Meet the State Board of Education’s Performance Target (90%).
2. Meet the Prescribed Graduation Rate Projection/Track individual target.
3. Be within two percentage points of the Prescribed Graduation Rate Projection/Track individual target and show overall improvement on the event dropout rate.
4. Develop a confidence interval approach in addition to the currently approved Graduation Rate calculations.

On July 28, 2008, ED approved the first three options. However, at the same time, the department denied the fourth option, stating,

Graduation rate is the measure of the number of students who graduated in the standard number of years with a regular diploma. It is a count, not based on a proportion or a sample. Thus, there is no inherent variability in this percentage that would require a confidence interval to ensure that the target has not been met, such as is the case with state assessment results.

Utah, which uses a grade 10 through 12 cohort to calculate graduation rates, submitted several related amendments this year (approved June 17, 2008). The state uses a three-year cohort because over one-half of its high schools do not include grade 9. Utah also lags reporting by one year so the Class of 2007 is the first graduating class for which a cohort calculation has been possible there. Utah will apply the following decision rules to its AYP calculations:

- If the cohort contains less than 40 students, there are too few students to make a reliable decision about the graduation rate, the graduation rate indicator is not applicable and so the unit “passes;”
- If the cohort is equal to or greater than 40 students and the graduation rate is greater than 85.7%, the unit “passes;”
- If the graduation rate is less than 85.7% but greater than the graduation rate for the prior year, the unit “passes;”
- In all other cases the unit “fails.”

Utah established its graduation target rate (85.7%) by using available data and estimating the graduation rate for the Class of 2007. That estimation, at the time that the cohort was in 10th grade, came out at just under 86%. According to state officials:
The figure of 85.7 also has a meaningful interpretation in terms of the event dropout rate. It can be obtained by accepting a maximum event dropout rate of 5 percent per year over the life of the cohort. Thus, 95 percent of the cohort remains at the end of 10th grade, 90.025 percent (95% of 95%) remains at the end of 11th grade, and 85.738 percent or (rounded) 85.7 percent (95% of 95% of 95%) remains at the end of 12th grade, the normative time of graduation.

This figure is not a starting point as such, but a constant applied every year, since under NCLB regulation 200.19(d)(1), "States are not required to increase the goals of its other academic indicators over the course of the timeline."

For the 2008 AYP reports only—if a student group makes safe harbor (with or without confidence interval) and if the n size is greater than 40 for both years and if the graduation rate is below 85.7% for a student group other than whole school, attendance for 2007 and 2008 will be used to determine if there is an increase in the additional indicator. This will only be used for 2008 to allow a baseline to be set for each student group for graduation rate.

**Growth Models**

Until early 2006, ED had not permitted states to use growth models in their educational accountability systems. ED’s position had been that growth models did not conform to the prescribed rules for AYP. In January of 2006, however, that changed and ED invited states to apply for participation in a Growth Model Pilot Program that would offer up to 10 states the option of using it for AYP. In a process that was described in the last two papers (Forte & Erpenbach, 2006; Erpenbach & Forte, 2007), nine states were eventually approved as part of ED’s pilot program. In late December 2007, ED announced that the growth model option would be available to any state provided that the state’s final assessment system was fully approved.

In 2007, five states sought approval to implement growth models (Colorado, Connecticut, Hawaii, Louisiana, and Pennsylvania). All but Colorado’s were denied either by ED or after peer review. Colorado’s request involved a growth model related solely to safe harbor determinations that was approved without going to peer review.

This year, 10 states submitted amendment requests pertaining to growth models. In the case of Delaware (approved June 5, 2008) and Iowa (approved July 28, 2008), both states simply affirmed their intent to continue previously approved growth models without changes as a part of their educational accountability systems. Hawaii included its growth model proposal in its accountability workbook amendment requests. In that state, ‘a student [would be counted] as ‘proficient’ (for accountability purposes) if the student’s growth rate on the Hawaii State Assessment, the Hawaii State Alternate Assessment, or the Hawaiian Aligned Portfolio Assessment places the student on a path toward the proficient benchmark for reading or math, or both, in three years.” ED denied the amendment on September 25, 2008, stating

Only states participating in the Department’s growth model pilot can include a measure of individual student growth in AYP determinations. . . . One of the criteria . . . to participate . . . is that the state

---

has a fully approved assessment system. . . . Hawaii’s assessment system . . . [is not approved] and, thus, the state is not eligible to participate in this pilot.

**Louisiana**, acknowledging that the state remains interested in the use of growth modeling in educational accountability determinations, requested approval of a “pilot growth model for use in our alternative schools.” Louisiana proposed the pilot approach because it was uncertain that a full model could be developed in time for use in 2007-08. According to the state’s proposal:

These [the alternative schools] are dropout prevention centers, career options programs, discipline centers, adjudicated facilities, and in a few instances, special education centers for our most severely handicapped students. Based on our 2007 testing population, these students comprise less than one hundredth of one percent (0.01%) of our students. Due to the transient nature of much of this population, even fewer meet the condition for inclusion in school accountability. We will continue to include them in district accountability, as we have done since 2003.

Most of the test results from alternative schools in our state route back to the students’ home schools, and we do not propose to change this. For those schools that are evaluated, we will continue to apply our approved accountability system with the growth model, providing an additional opportunity for these sites to achieve adequate yearly progress. . . . We envision a system similar to that approved for Delaware.

On September 24, 2008, ED denied the request. The rationale is interesting in that the department had previously approved the limited use of growth models for Colorado and Utah involving safe harbor reviews. The department’s decision with respect to Louisiana was similar to that given to New Mexico’s request earlier this year:

The Department cannot approve Louisiana’s request to establish a growth measure for use in its alternative schools. Only those states that are approved to participate in the Department’s growth model pilot program may include a growth model in their AYP determinations. To be approved to participate in the growth model pilot, a state must submit a growth model proposal to the Department for review by the Department and a panel of expert peers. Louisiana has not submitted a growth model proposal for peer review and has not been approved to participate in the growth model pilot.

**Missouri** (approved June 2008) and **North Dakota** stated that they had submitted growth model proposals to ED. 35 **Ohio** made changes to its minimum group size (uniform at 30), the last requirement necessary to implement a conditionally approved growth model. On June 12, 2008, ED denied North Dakota’s growth model proposal because it was not submitted in time for the Department’s April 2008 growth model peer review. ED further denied the state’s request that it conduct a “rolling review” of the growth model proposal but gave no specifics for the denial.

ED denied (June 17, 2008) **Oklahoma’s** request to use a growth model in its accountability system because the state included its proposal in its accountability workbook amendment proposals instead of submitting it independently to the department. **Pennsylvania** re-submitted its growth model proposal. According to the state:

35 ED has consistently used a separate, outside of the accountability workbook amendment process, submittal and review process for growth models. ED posts growth model information and decision letters at [http://www.ed.gov/lead/account/growthmodel/index.html](http://www.ed.gov/lead/account/growthmodel/index.html)
The essence of this proposal is that by using a projection metric, Pennsylvania will be able to assess the trajectory toward proficiency for each student, which results in a recalculation of proficiency at the district, school and subgroups levels. Pennsylvania’s proposed growth model keeps Pennsylvania on the trajectory for the NCLB target of 100% proficiency by 2014.

An August 18, 2008, decision letter from ED confirmed that the proposal had been denied. The department noted that: “The Pennsylvania growth model was forwarded to a group of peer reviewers through the Secretary’s growth model pilot. After considering the comments of the external peer reviewers, the Department did not approve Pennsylvania to implement its growth model in 2007-08 as noted in a letter to Pennsylvania on June 10, 2008.” As included in the discussion on growth models in Chapter 1 of this paper, Assistant Secretary Kerri Briggs wrote in the June 10 letter, “The heart of the peers’ concerns relates to the interaction of Pennsylvania’s existing performance index with the growth model. The Department has concerns about the appropriateness of allowing a state to include both a performance index and a growth model in its accountability system.”

Utah’s request was also denied by ED (June 17, 2008):

The Department declines to approve Utah’s request to include a growth model in its AYP determinations for 2007-08. Utah previously submitted a growth model proposal to the Department that was reviewed and determined not to meet the Department's core principles. As such, Utah was not approved to participate in the Department's growth model pilot program. Only those states that are approved for the growth model pilot may include a growth model in their AYP determinations.

Annual Measurable Objectives and Intermediate Goals
In their original accountability plans submitted in early 2003, states were required to describe the trajectory of annual measurable objectives (AMOs) and intermediate goals (IGs) that would lead from their baseline percent proficient to 100% proficient by the 2013-14 school year. IGs represent increases in the performance targets and can occur every year (in which case they would be equivalent to the AMOs), every other year, or every third year. Regardless of the frequency of these increases, the amount of every increase must be the same. Most states initially chose patterns of annual increases or an increase every third year, but a number later adopted a “back-loaded” model similar to one that Ohio originated. This model uses increases every third year in the beginning but accelerates to annual increases as 2014 approaches.36

This year, 10 states submitted accountability workbook amendments related to the AMOs and IGs. This compares to five in 2007, 14 in 2006, and eight in 2005. The sharp increase in 2006 was due in large part to the implementation of additional testing in grades 3 through 8. For the most part, ED has routinely approved requests in this area without extensive comment or conditions.

Georgia, which is transitioning to new mathematics achievement standards and assessments for grades 3-8, received approval (June 17, 2008) to adjust its AMOs and IGs for 2008 and 2009 only. The changes will be from 66.7 to 60.0 in 2008 and from 66.7 to 63.0 in 2009. The result is that for these two years, the increases, or intermediate goals, will not be equal to those in prior years or in future years (8.3%). In

36 According to a recent study, 23 states have “backloaded” their proficiency trajectories for achieving proficiency in reading or language arts and mathematics by 2013-14. In these cases, the States expected smaller achievement gains initially followed by steeper gains in later years. How and whether these states will meet their achievement targets without the law changing remains unclear (Chudowsky, N. & Chudowsky, V., 2008).
somewhat of a “first,” Idaho received approval (June 30, 2008) to “adjust the progression” of its AMOs by using the same targets in the 2008-09 school year that were used for the preceding two years. The approval was made because the state’s AMOs originally had all students reaching proficiency by 2012-13 instead of 2013-14. Idaho will push its remaining AMO targets back by one year following the 2008-09 school year.

Maine received approval (September 19, 2008) for a new set of AMOs. The revisions were necessitated by the state’s transition to a grade 3 through 8 assessment and the new Maine High School Assessment. In Mississippi, the state implemented “revised, more rigorous curriculum frameworks in language arts and mathematics, and new assessments that are aligned with the new frameworks.” Thus, Mississippi needed to set new AMOs and trajectories in conjunction with standard setting which was done during summer 2008. On September 11, 2008, ED approved the state’s revised AMOs together with a phase-in plan. North Carolina is administering a new reading assessment in grades 3 through 8, which will also require resetting the AMOs. The state is also considering whether to revise its mathematics AMOs at the high school level given the impact of new higher-level achievement standards. In a June 6, 2008, letter to the state, ED acknowledged the fact that North Carolina is in the process of amending its AMOs and cautioned that “before the revised AMOs may be used in AYP determinations, they must be submitted to the Department for review and approval.”

New Hampshire recently transitioned to a new high school assessment (New England Common Assessment Program). In late March the state submitted revised grade 11 starting points and AMOs (approved April 25, 2008). This year, New Jersey administered “redesigned” assessments in language arts literacy and mathematics in four (grades 5-8) of its seven state assessments. According to New Jersey, “The focus of these new assessments is to add additional rigor and alignment with our state’s [academic] content standards.” Administration of the new assessments will also be moved closer to the end of the school year and will require adjusting starting points, AMOs, and IGs. This process will also delay notification to school districts and holding AYP status and sanctions constant for one year (see related discussion under “Reporting” later in this chapter). In a July 16, 2008, decision letter, ED noted the pending adjustments stating, “New Jersey must submit the new AMOs for review and approval by the Department before they may be used in AYP determinations.”

Oklahoma indicated that the state will exclude pre-kindergarten students when calculating the OAI (attendance) for its elementary schools (approved June 17, 2008). In Rhode Island, the state decided to maintain its existing high school AMOs in spite of introducing new English language arts and mathematics assessments (approved June 11, 2008). Wyoming decided that it would “leave its current AMOs in place for 2007-08 and will revisit them in 2009-10.” The state is dropping its optional mid-year testing window in 2008-09.

Strategies to Enhance Reliability
States use a variety of strategies intended to enhance the reliability of their AYP decisions, and those have been chronicled in just about every paper in this series. Generally, the primary purpose is to reduce the likelihood that schools and districts will be misidentified as needing improvement when they do not. Every state is required to specify the minimum number of students required for AYP calculations and many also use confidence intervals around one or more of the AYP indicators as well. Some states also combine data across two or three years to help reduce the impact of year-to-year differences in student samples on AYP decisions.
Uniform Averaging
Provisions related to uniform averaging are found in section 1111(b)(2)(J) of the law. States may average data over two or three years when making AYP determinations. Although the statute refers to “uniform” averaging, ED has permitted “non-uniform” averaging since 2003. In the latter instance, states typically average data such that schools and districts may use scores from only the current school year or average across up to three years, whichever benefits the school or district more. This year, as last year, three states submitted a related amendment.

Oregon was approved (May 20, 2008) to:

Calculate AYP first based on two years of data. If two years of data does not provide a sufficient number of tests, ODE [Oregon Department of Education] will use up to four years of data. For extremely small schools, additional years of data may be included on a case by case basis as required to evaluate AYP. For schools and/or districts that have sufficient numbers of tests to be evaluated based on two years and do not meet AYP, ODE will calculate AYP based on the most recent year of data. However, in these circumstances, ODE will use the margin of error established for the two years of data to ensure that the rigor of AYP determinations is not reduced.

Like other states, Oregon will also make a safe harbor determination if a school or district does not make AYP using the above analyses. In the case of new schools, the state intends to make accountability decisions “as soon as sufficient data points are available.”

South Dakota developed new programming that permits averaging in making AYP calculations. On July 15, 2008, ED approved the state’s plan affirming that, “Under its uniform averaging procedure, two-year averaging will not be run unless a student group is equal to or greater than 10 across two consecutive years. In cases where there are 10 or more students in the current year, but not the previous year, only the current year’s data will be used to determine AYP.” Virginia has previously used uniform averaging procedures in its AYP calculations. However, the state has been implementing new assessments in grades 4, 6, and 7 as well as new end-of-course measures at the high school level. By the end of the current school year, the state will have three years of data in which to allow averaging in grades 3-8 and high school (approved June 24, 2008).

Minimum n
Until 2006, ED tended to approve states’ requests for changes to their minimum n’s for AYP purposes. Then, in 2006, at least 10 states requesting changes in their minimum n’s found that ED’s decision rules had changed. That change seemed to be triggered in large part by an April 2006 Associated Press (AP) article (Dizon, Feller, & Duncan, 2006, April 18) alleging that 1.9 million students nationwide were not being included in AYP determinations and that the primary reason was states’ minimum n’s (Forte & Erpenbach, 2006, pp. 4-5).38

37 The minimum n is the fewest number of students for which AYP will be calculated according to a state’s general AYP rules. If the number of students in a school is below the minimum n, the state must use another method for determining AYP. AYP does not have to be calculated for groups within a school (e.g., students with disabilities) that do not meet the minimum n. Some states have used different minimum n’s for different groups and for each of the AYP indicators but were required under new regulations to use uniform group sizes beginning in 2007-08.

Just prior to the publication of the AP article, the department had issued its NPRM related to alternate assessments based on “modified” student academic achievement standards (December 2005). This NPRM included a provision that would eventually end and rescind earlier ED approvals regarding the use of larger minimum n’s for student groups such as SWDs and limited English proficient students. The final regulations were released in April 2007 and included this provision. Beginning in 2007-08, states are no longer able to use a different minimum n across the student groups required for AYP determinations even if previously approved by ED, as set forth at §200.7(a)(2)(ii).

In somewhat of a small surprise, 10 states and the District of Columbia submitted educational accountability workbook amendments related to minimum group size this year. On July 28, 2008, ED denied the District of Columbia’s request to increase its uniform minimum n from 25 to 40. ED stated that, “The District of Columbia failed to provide evidence demonstrating that the approved minimum group size of 25 students does not yield valid and reliable results.” Massachusetts clarified its use and application of minimum group size when making AYP decisions (approved August 18, 2008). The state makes AYP determinations for schools and districts when:

1. (A) There are 40 or more subgroup members, AND (B) the number of subgroup members is at least five percent of students whose results are included in the school’s or district’s aggregate AYP calculation; OR
2. The number of subgroup members is 200 or more.

In its submittal for this year, Missouri sought to change its group size to 50 for all student groups. The state previously used 50 for LEP and SWDs student groups and 30 for other groups. According to Missouri, “The increase from 30 to 50 includes improvement in measurement precision and allows us to continue to provide more validity and reliability to decisions regarding AYP at the subgroup level.” On June 25, 2008, ED approved a minimum n change but limited the number to 30 for all student groups. Montana indicated that it intended to establish a minimum group size of 40, which would increase the all students group to 40 consistent with the number used for disaggregated groups. The state conditioned this request on ED deciding not to approve its small schools AYP model described earlier in this chapter. However, on July 16, 2008, ED denied the small schools amendment while approving a minimum n of 30—not the 40 requested by Montana.

In New Jersey, the state will use 30 for all student groups (approved July 16, 2008). They had previously used 20 for most groups and 35 for SWDs. The department’s decision letter included, for one of the first times, a caution that,

As proposed in the April 2008 notice of proposed rulemaking, the Department intends to use the National Technical Advisory Council, recently established by Secretary Spellings, to review all states’ minimum group sizes and provide guidance to the Department in order to ensure valid and reliable accountability systems that, to the maximum extent possible, include all students.

In Ohio, the state will use a minimum n of 30 for all AYP determinations (approved July 1, 2008). Ohio had previously used 45 for the SWDs student group, 40 for participation rate measures, and 30 for all other measures. Oklahoma will use a minimum n of 30 along with a confidence interval of 95% in making AYP determinations (approved June 17, 2008).

Oregon will move to a minimum group size of 42 but will make AYP determinations by combining two years of data. According to the state, “In cases where a school or district would otherwise meet AYP based on the most recent year of data, ODE will make an AYP determination based on a minimum of 21
tests in the most recent year.” South Carolina (approved June 5, 2008) will reduce the minimum group size for LEP students and SWDs from 50 to 40. The state will then use 40 for all student group AYP decisions. Washington will use a minimum n of “30 students (or 1 percent of the test population if the tested population exceeds 3,000 students)” for its AYP calculations (approved August 18, 2008). Wisconsin will reduce the SWDs group from 50 to 40, resulting in a uniform minimum n for all student groups (approved May 20, 2008).

Confidence Intervals and Standard Errors of Measurement
Most States now use confidence intervals39 (CIs) around their percent proficient indicators, with approvals granted through the amendment process. After early approvals for a few, a number of states were subsequently approved to increase the size of CIs from 95% to 99% and to apply CIs to any of the AYP indicators. A few states have also amended their plans to include the use of standard error of measurement (SEM) bands around their indicators. Until 2006, ED routinely approved most of these requests, with the only exception being a 75% maximum for safe harbor.

In 2006, ED was slow to respond officially to any new requests related to the use of confidence intervals other than those related to safe harbor; in most cases, ED’s posted decision letters made no reference to states’ requests in this area. In a letter to U. S. Representative Howard McKeon (Spellings, June 13, 2006), Secretary Spellings provided some insight into the reasons for the delays in responding to states’ requests: “When considering a State’s request for amendments to its accountability system that directly affect validity and reliability, the Department looks at the interaction of the full range of factors, such as the minimum group size, the confidence interval, if any, and the use of uniform averaging” (p. 3).

Since 2006, no state has requested an amendment related to confidence intervals or standard errors of measurement.

Safe Harbor
The term “safe harbor” is not actually used in the NCLB statutes and regulations. However, it is an apt description of provisions found under section 1111(b)(2)(I)(i) of the law. In its general form, safe harbor may be invoked when a student group misses the percent proficient target, but makes the other targets and reduces the percent of students scoring below the proficient achievement level by at least 10% from the previous year. The school or district makes AYP when it meets these safe harbor criteria.

This year, eight states sought related amendments compared to six last year and 10 in 2006. Perhaps the most interesting accountability amendment in this area last year was the one ED approved (May 23, 2007) for Colorado: “Colorado will add an additional safe harbor measure that uses a longitudinal model to compare the same students’ scores from the prior year to the current year.” The approval was for one year only with continuation contingent on the department’s review of impact data. In mid-May of this year, Colorado officials reported to CCSSO that ED had given verbal approval for an extension of the model, which they refer to as “matched safe harbor.” Approval was confirmed in ED’s August 4 decision letter although the department added, “Colorado will not apply a confidence interval to this application, nor will Colorado apply this additional measure of safe harbor to K-3 schools.” No rationale or explanation of these limits was provided. In making its request for continuation of this model, Colorado reported that only

39 CIs are statistical estimates of the range in which a school's or a district's “true” AYP score might fall given the score it actually obtained. For further information about confidence intervals, see Hill and DePascale (2003); Jaeger & Tucker (1998); Linn, Baker, & Betebenner (2002); or Marion, et al (2002).
17 additional schools (less than 1% of all schools in the state) made AYP than would have otherwise done so. Only two additional school districts were able to make AYP under the model.

In its February 15, 2007, letter to ED requesting initial approval of this model, Colorado provided important instructive background for other states considering adding this approach to their safe harbor review procedures.

After analyzing the current Safe Harbor results with the longitudinal methodology, we have found that more districts/schools/subgroups make the current Safe Harbor measurement than make the longitudinal measurement. It is truly more difficult to show a ten percent reduction of the percentage non-proficient for the same students than when you compare the cohorts. The only areas where the longitudinal reduction is more successful is at the school level, in reading for Asian students and at the district level, in reading for Asian students and English language learners.

We believe that there are two main reasons why the longitudinal reduction is a greater challenge than the current methodology. First, in math, the percentage of proficient students statewide declines as we advance from grade to grade. In third grade, 92.76% of students score proficient, but by the 10th grade there are only 66.33% proficient. When comparing third and fourth grade scores to fourth and fifth grade scores for example, statewide results show a decrease in proficiency. The percentage of students proficient in reading is fairly steady from third to tenth grade. Second, the data indicate that improving individual student growth is much more difficult than improving a school as a whole, over time.

According to ED’s 2007 decision letter:

Colorado has a data system that is able to track achievement as students progress through the school system. The longitudinal safe harbor will allow any district, school, or student group to make adequate yearly progress (AYP) provided there is a 10 percent decrease in the percentage of non-proficient scores in the current year compared to the previous year, comparing individual student assessment results in the current and previous year. Colorado will include this calculation for any district, school, or subgroup that does not meet the annual measurable objective or the traditional safe harbor, provided that the match rate for the group of students is 95 percent or higher. Colorado will not apply a confidence interval to this application.

Colorado appears to be the first state to draw on the provisions of section 1111(b)(3)(B): “Each State educational agency may incorporate the data from the assessments under this paragraph [those required for State assessment systems] into a State-developed longitudinal data system that links student test scores, length of enrollment, and graduation records over time.”

Colorado revised its alternate assessment in reading for 2007-08. Because of the changes, the test will not be comparable to the 2006-07 alternate assessment reading test. Therefore, the state proposed to “allow districts and schools to appeal the [alternate assessment] reading results from safe harbor calculations.” On August 4, 2008, ED denied the request, stating, “If a state is unable to compare the results of its tests to the previous year’s scores for whatever reason, it is not able to take advantage of the safe harbor calculation.” Kansas (approved June 6, 2008) “revised the calculation of 75 percent confidence interval that is applied to Safe Harbor calculations. The state will now use a standard error for the difference of proportions when calculating Safe Harbor.” Missouri submitted revised language to describe its safe harbor processes and related decision-making (approved June 25, 2008). These did not serve to substantively change or add to either the processes or resultant decisions. New Hampshire,
which implemented a new high school assessment this year, received approval (April 25, 2008) to “use a transitional measure to determine safe harbor by comparing the NHEIAP [old assessment] . . . results from May 2006 to NECAP [New England Common Assessment Program] results from October 2007 using an equi-percentile comparison.”

**Texas** proposed that, “If a district, charter, or campus meets the absolute standard on the other measure for all student groups that meet minimum size criteria (in addition to all students), they do not have to show improvement on the other measure as a condition of safe harbor.” The state argued that improving on the OAI linked to safe harbor reviews is an unreasonable requirement when schools or districts are already meeting, and often exceeding, the OAI target. ED, in a May 19, 2008, decision letter stated that Texas had “clarified that, in order to make Safe Harbor, a subgroup must decrease the percentage of students scoring non-proficient by at least 10 percent from the previous year and must either meet the target for the other academic indicator (attendance and graduation rate) or make improvement from the previous year.” **Wisconsin** clarified that science assessment results are used as a proxy in student group safe harbor calculations whenever attendance or graduation rates are not available. The state indicated that it would begin using disaggregated attendance and graduation rates after 2008. ED did not include a reference to this in its May 20, 2008, decision letter to the state.

Both **New Mexico** and **Utah** proposed a safe harbor growth model similar to that approved for Colorado last year. New Mexico’s proposal was subsequently denied and part of Utah’s was approved after being initially denied as well. In **New Mexico**’s case, ED noted in its July 23, 2008, decision letter that the state had previously submitted a growth model that was reviewed and determined not to meet the department’s core principles. ED concluded, “Only those states that are approved for the growth model pilot may include a growth model in their AYP determinations.” This decision stands in stark contrast to the approval given to Colorado (see above) last year. Like New Mexico, Colorado had previously submitted a growth model proposal that was not approved.

**Utah**’s request was denied on April 17, 2008. At that time, ED stated that Utah’s U-PASS (Utah Performance Assessment System for Students) did not include a requirement that all students be proficient in reading and mathematics by 2013-14. In response, Utah stated that it “is not willing to change its system so that all students are expected to score proficient in math and reading by 2014.” A Utah spokesperson stated that such is “not a realistic goal,” adding “It’s not an attainable goal right now.” ED denied a similar proposal in 2006 for the same reasons.

Utah’s request for approval included two parts:

For subgroups that do not meet the status AMO and do not meet the standard safe harbor requirement, a longitudinal student growth measure will be applied. The growth score is a longitudinal measure defined by comparing the same student from one year to the next year for language arts and/or math CRTs. A progress score will be determined for every student who is enrolled for a full academic year in the current school and who has a score (found anywhere in the state) for the previous year. The progress score is determined by the U-PASS progress tables. The individual progress score for each student will be averaged to determine the student group progress score. There are two progress tables; one each for grades 3-8 and grades 10-12. For language arts and math, grades 3-8, the student group must score at least 190 on the grades 3-8 progress table to make AYP. For grade 10 language arts, and grades 10-12 for math, the student group must score at least 180 on the high school progress table to make AYP. This is described on page 28 of the Workbook. Detailed information concerning the progress tables is in the attachment, “Longitudinal Approach Description” by Scott Marion, Center for Assessment.
If USED is unwilling to accept the growth calculation, USOE [Utah State Office of Education] requests a Safe Harbor Same Student Calculation: For subgroups that do not meet the status AMO and do not meet the standard safe harbor requirement, a Safe Harbor Same Student Calculation will be applied. This approach compares the performance of students in the student group enrolled for a full academic year that did not meet either the status or safe harbor targets with the performance of these same students in the prior year (as long as they were tested anywhere in Utah) to determine if there was a 10% reduction in the percentage of these matched students scoring below proficient.

On June 17, 2008, ED approved the second part of Utah’s request with the following conditions:

1. Utah will not apply a confidence interval;
2. K-3 schools will not be included in the calculations; and
3. Any district, school, or subgroup must have a 95 percent match rate in order to include the longitudinal safe harbor calculation.

Inclusion of All Students in Accountability
NCLB extended federal policies regarding the inclusion of students with disabilities (SWDs) and English language learners (ELLs; referred to as Limited English Proficient or LEP students in the law) in statewide assessment and accountability systems farther than any of the preceding versions of ESEA. SWDs and ELLs must now participate in assessments in all grades, 3 through 8 and once in grades 10 through 12, and their participation and performance must be reflected in the AYP analyses for every school, district, and state.

States continue to annually request a number of amendments related to the inclusion of SWDs and ELLs in their accountability plans. With respect to SWDs, most requests since 2005 have centered on the continuation of assessments based on “modified” student academic achievement standards. Amendments involving SWDs are described first, below. The description of amendments to inclusion policies for ELLs sought by states follows.

Inclusion of Students with Disabilities in Assessments and AYP
NCLB requirements for the full inclusion of SWDs in statewide assessment and accountability systems have posed a number of challenges to states ever since the law’s enactment. Like the past two years, states have again this year sought permission to exercise the “2% proxy” flexibility ED first offered for 2004-05 AYP decisions—to modify their definition of the SWDs student group for AYP decisions and to change AYP measures for this student group. Although regulations related to modified standards and assessments were promulgated in April 2007, ED has extended the related transition flexibility period. This is in recognition of the fact that the regulations are still very new and many states have not had sufficient time to develop or implement modified standards and assessments.

Modified Achievement Standards and the 2% Solution
On April 9, 2007, ED released final regulations related to the assessment of SWDs against “modified” student academic achievement standards (more commonly referred to as the “2% option”). The new regulations:

Provide States with additional flexibility regarding State, local educational agency (LEA), and school accountability for the achievement of a small group of students with disabilities whose progress is such that, even after receiving appropriate instruction, including special education and related services designed to address the students’ individual needs, the students’ individualized education
program (IEP) teams (IEP Teams) are reasonably certain that the students will not achieve grade-level proficiency within the year covered by the students’ IEPs (Federal Register, 2007, p. 17748).

Also included in the final regulations were provisions that:

- Require states to develop one or more alternate assessments for SWDs unable to participate in the regular assessments with or without accommodations as provided at §300.160(c). In addition to the assessments based on alternate achievement standards (1% assessments), states may offer assessments based on their academic content standards and “modified” achievement standards. The allowance for these alternate assessments based on modified achievement standards reflects a belief that there are some students with disabilities who are not able to demonstrate their grade-level content knowledge and skills on a state’s general assessments even with accommodations.

- Clarify that out-of-level assessments are not permitted for students assessed against modified achievement standards as found at §200.6(a)(3)(ii). The use of out-of-level assessments for students studying at lower academic levels than their grade level was prohibited under regulations issued by ED on July 5, 2002. In regulations issued on December 9, 2003, pertaining to significantly cognitively disabled students, ED recognized that “instructional” level (out-of-level) assessments “that meet the requirements of §200.1(d) may be considered to be an alternate assessment aligned with alternate achievement standards for the purposes of calculating AYP.”

- Permit the scores for students formerly receiving special education services to be included in SWDs student group AYP determinations for up to two years as provided at § 200.20(f)(2). A few states had argued that a longer period could be supported as long as students are still being monitored under IDEA provisions. The two-year limit may be a restriction that is contrary to IDEA if a formerly served student continues to be “monitored” for a longer period.

In 2004-05, more than 40 states applied to use the 2% option in their AYP calculations and most of these requests were approved. In 2006, at least 33 states applied for an extension of this option in making AYP determinations and 31 of these requested the same extension last year. For 2008, 25 of the 31 states re-submitted and three additional states submitted their first request in this area. A likely cause of the drop in the number of states re-submitting this year is the fact that ED has added the requirement of having a fully approved assessment system in place. Many states are still struggling to meet ED’s criteria for approval of their alternate assessments for significantly cognitively disabled students.

As noted on in the 2005 paper, “The continued interest in this option is somewhat surprising given that there is little evidence that it changes AYP outcomes to any appreciable degree and that States are prohibited from applying confidence intervals in these AYP determinations” (Erpenbach & Forte, 2005, p. 29). There were originally three options that states could choose from with respect to the application of modified achievement standards for SWDs participating in state assessments (see Spellings, 2005, May 10):

- **Option One**—States that do not have and do not wish to develop modified achievement standards (which must be based on grade level content expectation) may add a number equivalent to 2% of all students assessed to the percent of SWDs scoring at or above the proficient level. To qualify for this option, a state must describe how it will take six prescribed steps to build appropriate assessment tools for the target group of students.

- **Option Two**—States that have administered a high quality modified assessment statewide for two years or more . . . can use the results from this assessment for AYP purposes. Such assessments must be aligned with modified achievement standards as described above.
• **Option Three**—A state may offer another alternative for ED to consider.

However, in announcing the process that ED would follow with respect to educational accountability workbook amendments for 2008, Assistant Secretary Kerri Briggs stated in a letter to chief state school officers that Option 2 would no longer be available because it “permitted States to assess students on unapproved modified academic achievement standards . . . [an option] no longer available to States now that final regulations regarding those standards are in effect” (Briggs, 2007, November 19). A state wanting to exercise this option would first have to submit their modified academic achievement standards and alternate assessments for peer review and receive department approval.40

States that submitted requests to use Option One for 2007-08 AYP calculations were **Alabama** (approved June 10, 2008), **California** (denied September 26, 2008), **Delaware** (denied June 5, 2008), **Florida** (approved August 18, 2008, for reading/language arts only), **Georgia** (approved June 17, 2008), **Hawaii** (denied September 25, 2008), **Idaho** (approved June 30, 2008), **Illinois** (approved June 6, 2008), **Indiana** (approved July 15, 2008), **Iowa** (approved July 28, 2008), **Louisiana** (approved September 24, 2008), **Maine** (approved September 19, 2006), **Michigan** (approved July 15, 2008), **Minnesota** (approved August 27, 2008), **Montana** (approved July 16, 2008), **New Mexico** (approved July 23, 2008), **New York** (approved August 18, 2008), **North Carolina** (approved June 6, 2008), **Pennsylvania** (approved August 18, 2008), **South Carolina** (approved June 5, 2008), **Tennessee** (approved July 28, 2008), **Virginia** (approved June 24, 2008), and **Washington** (approved August 18, 2008).

**Alaska** decided not to seek an extension (approved February 12, 2008) of its prior approval under Option One stating, “At this time, Alaska is analyzing what other states are doing regarding assessments relating to modified [student] academic achievement standards, and studying what effect implementation might have in Alaska.” **New Hampshire** stated that it did not seek an extension because its assessment system is currently in approval pending status and therefore not eligible for additional flexibility.

While ED denied **California’s** continued use of “automatic increased proficiency levels” for SWDs under Option One (see above) the department did approve (September 26, 2008) the inclusion in 2007-08 AYP determinations “the scores of students with disabilities in grades 3-5 who are proficient (up to a 2.0 percent cap at the district and state levels) on the CMA [California Modified Assessment], an alternate assessment based on modified academic achievement standards.” The assessment will still be subject to approval through the peer review process. ED acknowledged in its decision letter that the State Board of Education will not formally adopt the new modified achievement standards until November of this year. The department required that California re-compute all of its 2007-08 AYP calculations for schools and districts with CMA-takers. According to California officials, some percent proficient results will go up and some will go down. ED’s decision letter stipulated that all schools and districts would carry forward into 2008-09 their 2007-08 AYP status and also set forth how changes, if any, resulting from the recalculations based on CMA-takers would be handled.

---

40 In March 2008, ED conducted its first round of peer reviews related to modified achievement standards and assessments. Six states (**Kansas**, **Louisiana**, **Maryland**, **North Carolina**, **North Dakota**, and **Oklahoma**) submitted peer review “packages.” None of the six “passed.” All six states were invited to a meeting with ED officials in May to “discuss the work needed to complete the peer review process.” The meeting resulted in a “lessons learned” paper (Filbin, 2008) that appears to have introduced additional criteria that will be applied to subsequent peer reviews of states’ assessments based on modified achievement standards. The six states received formal letters in early September detailing the additional “work needed to complete the peer review process.”
In denying Delaware’s proposal to continue use of the “proxy method,” or Option One, ED pointed out that “to be eligible for the transition flexibility, a state must show that it is making substantial progress in developing an alternate assessment based on modified academic achievement standards (AA-MAS).” Delaware had sought bids for an alternate assessment based on modified standards last year but none were accepted due to budget constraints. ED concluded that the state was "not able to provide evidence that an AA-MAS will be developed and implemented. In Hawaii’s case, ED noted that state’s assessment system is not yet approved and thus, it is not eligible for this flexibility.

Kansas (approved June 6, 2008), North Dakota (approved June 12, 2008), Oklahoma (approved June 17, 2008), and Texas (approved May 19, 2008) are four states that indicated they would follow Option Two by submitting their modified achievement standards and alternate assessments based on those standards for peer review and department approval.

Colorado requested interim flexibility to continue using a model approved in each of the past three years. The model is applied to districts and schools appealing AYP determinations when the only reason for the identification is failure to meet current performance targets for SWDs and the student group met 2004-05 targets. Colorado acknowledged in its request that states seeking interim approval are expected to “be well on the way to a modified assessment system.” However, based on extensive research, Colorado noted that it would not pursue development of a modified assessment system and instead would focus on a different approach, one it argued will require more of students and schools. On August 4, 2008, ED denied the request, stating simply that, "only states that are moving expeditiously to administer alternate assessments based on modified academic achievement standards are eligible for this flexibility." Acknowledging that the state did not intend to do this, ED concluded, "As a result, Colorado is not eligible to use the transition flexibility."

According to Colorado officials:

After bringing together various stakeholders and looking at multiple years of data, there has been consensus in the state that an additional layer of standards and a different assessment is not what is best for our students. We have very few students who are not showing some movement in CSAP [Colorado Student Assessment Program], and a few students who are topping out of CSAPA (the alternate assessment). However, CDE believes that the work we have done and plan to do, with regard to data analysis and the appropriate use of accommodations, should warrant additional interim flexibility.

We determined that students who were four standard deviations below the CSAP mean would be considered students in the gap between assessments. In 2005, there were 1,737 students in grades 3-10 who met this criterion for reading (0.038% of the tested population). In 2007, there were 1,653 students (0.035%). If students remain in this low level of CSAP performance, we’d consider them to be in the gap. However, only 212 students who were outliers in 2005 were again outliers in 2006. These results follow the same pattern for 2006 and 2007. Clearly, there are very few students in Colorado that remain as outliers. When our stakeholders look at this [sic] data, they conclude that if students are able to access CSAP over time, then we should continue to provide them with the opportunity to meet high standards and not create a modified assessment system.

Additionally, for the CSAP, we are focusing on providing appropriate accommodations for students with disabilities. We looked at students with disabilities who score in the lowest third of the Unsatisfactory category for three years (another way to possibly identify students who might be considered in the gap). Of the 1,192 students in grades 3-10 in this situation, only 525 students
received accommodations on the CSAP for all three years, and 249 students never received any accommodations. Only 62.3% of students with disabilities received accommodations in 2007 on the reading CSAP. As time in school increases, the student’s likelihood of receiving accommodations decreases—at elementary school 70.2% do, while at high school, only 48.7% receive accommodations. In response to this data analysis, the Colorado Department of Education created the “Colorado Accommodations Manual” posted at www.cde.state.co.us/cdeassess/documents/csap/manuals/2007/CO_20ACCOMM_MANUAL_1017_2007.pdf to provide additional guidance to the state on determining and providing appropriate accommodations for students. After a few years of focused attention on accommodations, the state will review the data to reflect on the “fit” of assessments for students. At that point in time, if it appears clear that additional standards and assessments are what is best for students, the state will move forward in that direction. But for the time being there is no reason to create modified standards and expectations in Colorado until we are convinced that there are some students who cannot meet the current standards. The Colorado Department of Education wants to be certain that students are given every chance and opportunity to show that they can meet the state content standards before we change our expectations for them.

**Maryland** requested continuation (approved June 17, 2008) of its modified achievement standards appeal process model first approved in 2006 (Option Three) for students in grades 3-8 and high school students who took end-of-course assessments in summer 2007, October 2007, and January 2008.41 Students in grades 3-8 will not take the state’s new modified assessment until spring 2009, although the new modified assessment was given to high school students in May 2008.

Under the appeals process model, schools that do not make AYP based solely on the SWDs group have IEP teams review individual student IEPs to affirm the identity of those students who might have received proficient scores on a modified assessment if one had been available. In its 2007-08 accountability workbook submittal, the state reported that the “process generated appeals from 177 schools [last year].” As a result of the appeal process, an additional 96 schools were identified as having made AYP. In 2006, the state received appeals under this option from over 160 schools representing approximately 1,200 students. Maryland approved about 66% of the reading appeals and about 50% of the mathematics appeals, resulting in 115 schools that made AYP.

**Massachusetts** received approval (August 18, 2008) to continue the approach approved annually since 2005 with respect to the assessment of students against modified achievement standards. The state’s approval is identical to last year’s and involves assigning “100 ‘performance index’ points to students selected based upon set criteria equivalent to 2.0 percent of all students assessed. For any school or district that did not make AYP solely due to its students with disabilities subgroup . . . this adjusted index score [will be used] to re-examine if the school or district made AYP for the 2007-08 school year.” Massachusetts’ request for a one-year extension was accompanied by considerable documentation of the state’s planning related to the development of an alternate assessment based on modified student academic achievement standards.

---

41 Maryland’s assessment system received full approval in June 2006. Modified student academic achievement standards were included in the approval. Maryland is awaiting peer review results of its modified assessments for grades 3-8 and high school.
**Defining the Students with Disabilities Student Group for AYP Decisions**

Until 2007, only two states had been approved to extend the definition of the SWDs group to include exited students. **Georgia** was approved in 2003 to include in the SWDs group those students who were still receiving special education services either in the form of monitoring or support in the transition to the regular classroom. **South Carolina** was approved in 2004 to include in the SWDs student group those who no longer are directly served in special education as long as their IEP called for monitoring services. Since then—and until last year—although several states made similar requests, ED apparently denied these, usually without written comment.

However, the new regulations adopted in April 2007 provided that states could include formerly served SWDs for up to two years in making AYP determinations for this student group. Last year, five states (**California**, **Idaho**, **Iowa**, **Pennsylvania**, and **Utah**) submitted requests in this area and all were approved. On the other hand, ED denied the requests of two states (**Georgia** and **Wyoming**) concerning inclusion requirements for AYP determinations. In each case, the effect of the states’ requests would have served to exclude some students receiving special education services from AYP calculations for that student group.

This year, 15 states submitted educational accountability workbook amendments pertaining to composition of the SWDs student group. **Alaska** (approved February 12, 2008) will include “students who were formerly classified as students with disabilities in that subgroup for two years after they have exited special education.” Conversely, **Iowa** will no longer exercise this option in its AYP determinations (approved July 28, 2008). The state reported that it has been unable to consistently validate the status of students in this group in time to meet parent notice requirements prior to the beginning of the next school year. **Colorado**, through “its appeals process will include former students with disabilities in the students with disabilities subgroup for two additional years after they are no longer receiving services” (approved August 4, 2008).

**Connecticut** (approved July 15, 2008), **Hawaii** (approved September 25, 2008), **Indiana** (approved July 15, 2008), **Kansas**, **Maryland**, **New Mexico** (approved July 23, 2008), **Nevada** (approved June 10, 2008), **Rhode Island** (approved June 11, 2008), **South Dakota** (approved July 15, 2008), **Wisconsin** (approved May 20, 2008), and **Wyoming** (approved July 15, 2008) will also include students who have exited special education services for up to two years in AYP decisions. **Missouri** will include previously served SWDs for up to two years on appeal by school districts (approved June 25, 2008).

ED’s approval of **Hawaii**’s amendment included a stipulation not seen in the approvals for any other states with respect to the inclusion of formerly served in the SWDs student group AYP calculations. ED stated:

> Please note that this approval only relates to the proficiency calculation for the students with disabilities group. As noted in the Department’s regulations and non-regulatory guidance [not cited; guidance has no regulatory standing], Hawaii is not required to include former students with

---

42 ED included in its April 6, 2007, regulations pertaining to modified academic achievement standards language at §200.20(t)(2) indicating that the scores of formerly served SWDs could be included for up to two AYP determination cycles. However, it is only in the department’s discussion of “Other Provisions Addressed in These Regulations” (p. 17750), that it is specifically provided that states “would not be able to include the scores of former students with disabilities as part of the students with disabilities subgroup in reporting any other information (e.g., participation rates) under Title I.” The extent to which this may be an issue for other states is unknown.
disabilities when determining whether the school or district meets the state’s minimum subgroup size but it also may not include former students with disabilities when calculating the participation rate for the students with disabilities group. The flexibility to include former students with disabilities who no longer receive special education services for two additional years relates only to assessment scores that are used to calculate proficiency rates. Hawaii must correct this for the AYP determinations based on assessments administered in the 2008-09 school year.

**Changing AYP Measures for the SWDs Student Group**

As summarized earlier in this chapter, Georgia re-submitted two related measures that were both denied by ED last year. The first involved the definition of FAY for SWDs: Students reported to be receiving special education services from the fall enrollment count (first Tuesday in October) through the spring testing window. The state argued that this ensures that the group reflects the achievement of “children served by special education programs and services during the school year rather than students who might have only received services for a matter of days.” On June 17, 2008, ED denied the request, stating, “The length of time a student receives services has no impact on whether a student meets the definition.”

Georgia’s second request involved recognizing special education diplomas as regular diplomas for AYP measures. ED has consistently denied such requests and the department again denied it on June 17, 2008. ED cited section 1111(b)(2)(C)(vi) as the basis for the denial: “A regular diploma must be aligned to regular grade-level content standards, as defined by the state.” Hawaii reports SWDs who are in “out-of-level” grade placements or returned high school seniors who did not pursue a high school diploma as being in “Grade 31” status. Hawaii confirmed that SWDs in “Grade 31” status must take the state assessment documented in their IEPs, i.e., the age-appropriate regular or alternate assessment. ED approved the plan (September 25, 2008) and clarified that “If the student is placed out of level, the student will take the assessment for the grade in which the student should be enrolled.”

Louisiana has not yet received approval for its alternate assessment for significantly cognitively disabled students. Although the state recently administered a new alternate assessment, it will not have completed standard setting processes in time to include the results in 2007-08 AYP determinations. Therefore, Louisiana sought approval to not include the results in 2007-08 AYP measures. Instead, a zero would be assigned “to tests administered to students who are eligible, regardless of the 1% cap.” Students would still be required to participate in the alternate assessment. On September 24, 2008, ED denied the request stating only that “All students in the grades assessed must be included in annual AYP determinations.” However, the department also required that preliminary AYP results be available not later than October 31, 2008, and prescribed a transitional plan identical to that set forth for Texas; another state implementing a new alternate assessment (see the section on “reporting” later in this chapter).

Maine clarified that, beginning in 2008-09, students “with the most significant cognitive disabilities may take the alternate assessment based on alternate academic achievement standards in one subject and the general assessment in the other as determined by the student’s IEP Team” (approved September 19, 2008). Massachusetts sought approval to increase from 1.0 to 1.5 the statewide cap on the percentage of proficient or above scores for students with the most significant cognitive disabilities taking alternate assessments based on alternate achievement standards that could be included in AYP calculations. ED denied the request on August 18, 2008, citing §20013(c)(4) which “explicitly prohibits a state from exceeding the cap of 1.0 percent.” Missouri clarified the manner in which SWDs participate in its statewide assessments and how their scores are included for AYP determinations (approved June 25, 2008).
Montana confirmed that it will continue to allow every small district (i.e., those with fewer than 200 students in the tested grades) to count up to two proficient scores based on student achievement on alternate assessments aligned to alternate achievement standards when making AYP determinations regardless of whether this exceeds the 1% cap (originally approved by ED in 2005 with annual extensions). The state, which has approximately 50% of its school districts with total enrollments of fewer than 100 students, reported that “It has been our experience that the granting of an exception for districts with fewer than 200 students has not resulted in the state exceeding the statewide 1% cap.”

On July 15, 2008, ED approved West Virginia’s request to establish a process for school districts:

To request a waiver to exceed the 1.0 percent cap on the percentage of proficient or advanced scores of students with the most significant cognitive disabilities who take West Virginia’s alternate assessment based on alternate academic achievement standards that may be included in adequate yearly progress (AYP) determinations. In granting waivers, however, the state as a whole may not exceed 1.0 percent of all students in the grades assessed.

Inclusion of English Language Learners in Assessments and AYP
In previous papers on states’ annual requests for amendments to their accountability workbooks, it has been observed that prior to the enactment of NCLB a number of states excluded many English language learners (ELLs) from participation in statewide assessment and school and district accountability decisions. Thus, when NCLB was signed into law and included significant provisions for the inclusion of these students in state’s assessment and accountability systems, few states were prepared to meet the requirements. Additionally, states were required under Title III to conduct annual measures of the English language acquisition of these students.

Less than one year after the initial approval of states’ accountability systems in 2003, ED offered some flexibility for the inclusion of ELLs through a communiqué from then Secretary Rod Paige (Paige, 2004, February 20). The Secretary’s new flexibility allowed states to:

- Exempt ELLs from the reading or language arts assessments during their first year of enrollment in U. S. schools;
- Exclude reading or language arts and mathematics scores for these same students from AYP calculations that year;
- Count these same students as participants in the reading or language arts assessments if they take the English language proficiency assessments that year; and,
- Extend the definition of ELL to include students who had exited ESL or bilingual education programs within the previous two years.

Since the 2004 communiqué from Secretary Paige, nearly all states had taken advantage of these options by 2004-05. Thus, ED’s subsequent release of regulations providing for this same flexibility on September 13, 2006, did not have much effect on states’ AYP models. However, as noted below, ED now appears to be taking steps to restrict the elements that can be used to calculate AYP when including exited ELLs in student group determinations.
**Inclusion Policies Generally**

**Modify the Definition of Newly Arrived**

**Alabama** (approved June 10, 2008) will change its definition of newly arrived LEP students from those in their “first academic year of enrollment” to those in their “first twelve months” in U.S. schools. **Colorado** clarified (approved August 4, 2008) that:

Students with limited English proficiency can be classified as not English proficient (NEP), limited English proficient (LEP), or fully English proficient (FEP). All groups are included in the LEP subgroup for accountability purposes. To exit the LEP subgroup for accountability purposes, a student must be proficient in all modalities of English, including reading, writing, comprehension, and speaking, as measured by the Colorado English Language Assessment. In addition, the student must score proficient or above on the state’s reading/language arts assessment.

Earlier in 2008, ED Title I monitors had determined that Colorado was permitting exemptions from testing on the basis of language for some LEP students. Monitors had made a similar determination during a 2005 visit to the state. As a result, ED notified Colorado that it could not make such exemptions and would lose Title I funds if corrective action was not taken immediately. Reporting on the outcome of the latest Title I monitoring in Colorado, Brownstein (2008, September, p. 6) observed that although the state agreed to make the required changes in its policies for inclusion of LEP students in its assessments,

The state noted that . . . NCLB repeatedly (38 times) cites the need for “valid” and “reliable” assessments. Respected researchers agree that students who are limited in English proficiency may not demonstrate their reading content knowledge when tested in English. . . . An assessment in English will create an invalid reading test score for accountability purposes for students whose language proficiency is a barrier to assessing the document.

Brownstein also noted that a state representative acknowledged that there was “a level of civil disobedience in our pushback” but “Colorado will continue to press its arguments in a different forum, perhaps ED’s LEP Partnership . . . because we have difficulty accepting this from a fairness and validity standpoint.” Similar objections have been raised in Michigan, New York, Virginia, and Washington.

**Connecticut** received approval (July 15, 2008) for its “three-step process” designed to help school districts determine whether a student can be considered as an English language learner. The approved process:

1. Administer a three-question survey; 2. Use a dual language test or an observation procedure to determine the student’s dominant language; and 3. Determine limited English language proficiency based on three indicators:
   - a standardized English proficiency test;
   - an oral interview conducted in English; and
   - an examination of other indicators, such as report cards and test scores that demonstrate whether achievement in academic subjects taught in English is below average.

The **District of Columbia** (approved July 28, 2008) clarified that an LEP student “is a linguistically and culturally diverse student with an English language proficiency level that does not allow the student to participate in the general program of school without alternative language services.” Students are exited from the LEP student group when they attain “fluency in English language proficiency as measured by achieving a proficiency level of 5.0 or above on the ACCESS for English language learners test.” **New**
Mexico clarified that it will define English language learners in that state consistent with the provisions of section 9101(25) of NCLB statutes (approved July 23, 2008).

**Delay the Time Before Assessment Participation Is Required**

**Pennsylvania** re-submitted a proposal that could potentially delay the time at which the state would require ELLs to participate in its reading assessments. ED earlier denied (July 2, 2007) a proposal in which the state had sought to delay their participation until the students had scored in level three (of five) or higher on the state’s English language proficiency assessment or had reached their third year of receiving English as a Second Language (ESL) services. ED stated that it “cannot approve this request because it is not consistent with the statute or in line with the Department's September 2006 regulations regarding ‘recently arrived’ LEP students, who may be exempted from one, and only one, administration of the reading/language arts assessment in their first 12 months of schooling in the United States.” Pennsylvania continues to argue that, “Students who are just beginning to learn the letters and beginning concepts of print in English will have a very difficult and frustrating experience with the PSSA Reading. It is almost guaranteed that they will not be able to make any meaningful sense of the written text.” On August 18, 2008, ED again denied the proposal with an almost verbatim response.

**South Carolina** (approved June 5, 2008) and **Wisconsin** (approved May 20, 2008) will exempt ELLs from one administration of their reading assessments during their first 12 months attending U. S. schools.

**Washington** also re-submitted a related amendment request that ED had denied on June 4, 2007. The state again sought to exempt LEP students in their first year of enrollment in a U. S. school from having to take any assessment in which all of the items require reading English. The state also wanted to not count, for up to three years, results for LEP students taking assessments in which all of the items require reading English (unless the students were able to demonstrate at least an intermediate proficiency in English). On August 18, 2008, ED again denied the amendment in a virtually identical manner:

> Washington’s request exceeds the flexibility afforded by these regulations [§200.20(f)(1)(iii)] and conflicts with the statutory requirement to include LEP students in a state’s assessments under section 1111(b)(3)(C)(ix) of the ESEA. Additionally, Washington may permit any language translation or accommodation that may be necessary, so long as it does not invalidate the test results, but the students must be assessed.

**AYP Policies Related to the ELL Student Group**

As noted in the 2005 paper (Erpenbach & Forte, p. 32),

> ED has consistently allowed states to extend their definition of ELLs for AYP purposes to include students who exited ESL or bilingual education programs within the previous two years. NCLB provisions at section 9101(25) provide so broad a definition of the term, “limited English proficient,” that it would seem possible for states to include formerly served ELLs in group AYP decisions for much longer than two years.

---

43 Reporting in *Education Week*, Mary Ann Zehr chronicles the educational challenges presented by a Hmong 15-year-old who arrived in a California high school at the age of 15 having never attended school and speaking no English. Four years later, the student was a senior, and although she had made solid academic progress, she had not been able to score as proficient in the high school exams used for AYP. According to Zehr, “Her experience is an example of why some educators . . . here say the accountability provisions of the law don’t provide a complete picture of the quality of education at a school that has a high number of ELL students” (2008, June 4, pp. 1 and 11).
Nevertheless, ED has maintained—and codified in September 2006 regulations—that the limit for states to continue to include formerly served, or exited from services, ELLs in that student group for AYP measures is for up to two years after they no longer meet the State’s definition for LEP. The key appears to be how a state determines the point at which a student no longer meets its definition of limited English proficient (see section 9101(25) for the federal definition of LEP).

A new issue that arose this year for at least a small number of states was the matter of exactly what elements (participation rate, achievement scores, or other academic indicators) can be included in calculations when formerly served or exited LEP students were included in that student group for accountability measures. The new issue was triggered when at least one state was told by ED officials that they could use only achievement scores for formerly served LEP students when including them in that subgroup for AYP determinations (the state uses all three elements and has done so since 2004). The ED official cited non-regulatory guidance issued by the department in May 2007. That guidance (which is exactly what it is, “non-regulatory”) includes at pages 11-12 a question and answer that taken literally would restrict a state to using solely achievement score data for formerly served LEP students.

Former Secretary Rod Paige’s February 20, 2004, key policy letter provided states the initial flexibility for this inclusion simply stating that LEP students who had attained English proficiency could be included for up to two years “in measuring AYP” (without qualification of any kind). Neither any of the subsequent key policy letters from ED on participation rates or LEP students nor the September 2006 regulations and April 2007 revisions included any mention of delimiting the AYP elements that a state can use when including formerly served LEP students in subgroup measures. It would appear that the inclusion of the limitations in the May 2007 non-regulatory guidance are based on an “interpretation” of the September 2006 regulations since there is no other applicable statutory basis. The non-regulatory guidance interpretation followed more than three years after Secretary Paige’s key policy letter during which time almost every state had incorporated the flexibility provided in its accountability workbook. The extent to which ED decides to pursue this issue with other states remains to be seen.

In an Education Week article earlier this year, Zehr (2008, March 26. p. 8) reported on efforts of states to begin to move toward increased use of the results of English language proficiency assessments to decide when ELLs no longer need specialized instruction and may exit those services. Zehr found considerable differences among states in terms of the degrees of flexibility they afford schools and districts in the determination of when ELLs no longer need services. She also noted that, “The most widely used English-language proficiency test is the ACCESS for ELLs developed by the World-Class Instructional Design and Assessment consortium, or WIDA.” According to Zehr, WIDA representatives recommend that states let school districts “use a number of measures to decide whether ELLs no longer need specialized services.”

Inclusion in Statewide Assessments, AYP, and English Language Acquisition Measures

California, “In determining whether or not the LEP subgroup is numerically significant . . . will add in Reclassified-Fluent-English-Proficient (RFEP) students if their results are included in determining whether or not the LEP subgroup met the annual proficiency goals” (approved September 26, 2008). The state reported that it is unable to validate the exit status of either SWDs or ELLs through its student information system. Hawaii will continue to use an off-the-shelf version of the LAS Links assessment to measure the English language acquisition skills of LEP students (approved September 25, 2008). This assessment is not aligned to the state’s English Language Proficiency Standards. Hawaii argued that it needed interim approval of this approach while it develops an English language proficiency assessment and related procedures. In its approval, ED cautioned that “approval of this amendment does not constitute approval
of the English language proficiency assessment, which must satisfy the requirements of Title III of the ESEA.”

**Kentucky** (approved May 20, 2008) clarified that recently arrived LEP students “will be given a one-time exemption from the reading/language arts assessment.” The state also clarified that, “LEP students in their first year of enrollment in U. S. schools [will be required] to participate in the state-required science assessments.” **Maine** will exempt recently arrived LEP students from one administration of Maine’s reading assessment and will not count their reading or mathematics scores for AYP determinations for that single year (approved September 19, 2008).

**Massachusetts** wanted to change its method of “assigning reading/language arts index points to certain . . . LEP students.” The state submitted impact data in support of its plan to assign index points for “LEP students who are in their second year of U. S. schooling and have not yet demonstrated English language proficiency at the level required to perform Massachusetts Comprehensive Assessment System (MCAS) tasks in English.” Index points are assigned according to a student’s progress toward achieving English language proficiency as measured by the Massachusetts English Language Proficiency Assessment. Index points are then incorporated into the state’s Composite Performance Index which is used to make AYP determinations. The maximum number of points awarded per student would be 75, below the 100 awarded for a student demonstrating proficient performance on the MCAS. ED denied the plan on August 18, 2008, stating, “Neither the statute nor the regulations regarding the inclusion of LEP students in State assessment and accountability systems provides the flexibility to permit Massachusetts to use an English language proficiency assessment for AYP determinations. Doing so would mean that LEP students would not be held to grade-level content standards.”

**Missouri** submitted amendments clarifying the manner in which ELLs are included in the state’s English language proficiency assessments and regular assessments (approved June 25, 2008). The amendment ensures full compliance with NCLB requirements. **Nebraska** received approval (September 3, 2008) to permit school districts there to “allow districts, in making AYP determinations for the LEP subgroup, to include, for an additional two years, former LEP students who have exited from LEP services or from the district’s LEP program.” **New Hampshire** clarified that assessment results for LEP students who have been enrolled in U. S. schools for less than 10 months will not be included in the state’s AYP calculations (approved April 25, 2008).

**Virginia** sought to increase the 2009 target for Annual Measurable Achievement Objectives (AMAOs are required under Title III of NCLB) related to the percent of ELLs making progress and attaining English language proficiency. The target for those making progress would increase from 40 to 45% and for those attaining proficiency from 30 to 35%. Virginia also reported that it would set AMAOs for future years after data from its newly adopted English language proficiency assessment have been analyzed. However, ED denied the request on June 24, 2008, stating that, “Virginia must have this requested change processed and approved through the Office of English Language Acquisition (OELA).”

**Inclusion of Formerly Served, or Exited, Students in Student Group AYP Calculations**
Four states submitted related requests in 2007, and nine states submitted a related amendment this year. In **Arkansas**, ED approved (June 26, 2008) the state’s clarification with respect to criteria for exiting ELL services. The criteria were determined to be consistent with those prescribed by ED’s Office of English Language Acquisition.
On July 15, 2008, ED approved Connecticut's proposed criteria for ELL students to exit that status:

1. Score at the highest level on the LAS Links English proficiency assessment; and
2. Meet or exceed the Connecticut Mastery Test (CMT) "proficient" level for reading and mathematics; and
3. Meet the "basic" level on the CMT for writing.

Iowa rescinded its practice of including formerly served LEP students in student group AYP calculation (approved July 28, 2008). The state reported that it has been unable to consistently validate the status of students in this group in time to meet parent notice requirements prior to the beginning of the next school year. Kansas (approved June 6, 2008) will “continue to include in the LEP group students who were formerly LEP students for an additional two years after they no longer meet the state’s definition of LEP.” Maine will include formerly served LEP students in student group AYP determinations for up to two years after exiting services (approved September 19, 2008).

Missouri clarified the state’s criteria for exiting ELL services but affirmed that formerly served students would be monitored for two years and included in the student group’s AYP determinations (approved June 25, 2008). Nebraska (approved September 3, 2008) and New Hampshire will include students who have exited LEP status as part of that student group in both assessment and accountability reporting for two years (approved April 25, 2008). New Hampshire defines “exit status” as “The student has earned: 1) a Composite Proficiency Score of 5.0 on the most recent ACCESS for ELLs® test, and 2) scores at or above 4.0 in all the proficiency sub-domains, to include listening, speaking, reading, and writing.”

ED approved (June 17, 2008) Utah’s clarification of the state’s definition of LEP and exit criteria from that student group. Exiting ELL services in Utah is based on attaining English language proficiency (ELP) as determined by the state's ELP assessment and proficiency on the English/language arts tests. However, in its decision letter, ED cautioned that Utah may have to make some changes under the pending Notice of Proposed Interpretations published on May 2, 2008.

West Virginia (approved July 15, 2008) will exit students from the LEP student group whenever

a. A student no longer meets the definition of LEP and no longer participates in alternative language programs or receives monitoring services; and
b. A student scores above level five on the West Virginia Test of English Language Learning (WESTELL) for two consecutive years or tests proficient for two consecutive years on the West Virginia Alternate Performance Task Assessment (APTA); and

c. A student scores at mastery level or above on the West Virginia Educational Standards Test (WESTEST), Reading Language Arts Assessment (grades 3-8 and 10) or Reading Language Arts end-of-course exams (grades 9 and 11), or the student scores at mastery level or above on the APTA, alternate assessment based on alternate academic achievement standards.

LEP Student Group Safe Harbor Review

Georgia requested approval to implement a “second look” at ELLs when they fail to meet AMO targets on state assessments in mathematics or English language arts/reading. In this “second look,” Georgia planned to recognize as proficient on statewide assessments (if they fail to meet targets in those assessments) ELLs who were enrolled in U. S. schools for less than three years and performed at the lowest two performance bands on its English language proficiency assessment if these students moved positively from one performance band to another. ED denied the request (June 17, 2008), stating that it was “not covered by these [the narrow exceptions related to exempting recently arrived LEP students]
exceptions nor is it consistent with the statute and regulations requiring AYP determinations to be based on the student's performance on mathematics and reading/language arts content assessments."

**Use of Alternate Assessments for ELLs**

Last year, seven states submitted amendments or clarifications related to changes in their use of alternate assessments for ELLs. This year, no states submitted related requests. Last year's requests resulted in large part from a rejection of many previously developed alternates during the federal standards and assessments peer review process. The lone exception was an approval (tentatively on January 31, 2007, and finally on July 16, 2007) of an alternate assessment for some ELLS in Virginia. The model approved in Virginia appears to be working very well. According to an October 2, 2008, article in the *Washington Post*, "This year, Virginia for the first time rated thousands of English learners, many of them from immigrant families, through samples of schoolwork instead of a regular test. That led scores to jump" (Glod & de Vise, 2008). According to a spokesperson at the Virginia Department of Education:

Beginning with the 2006-2007 school year, Virginia requested (December 16, 2006) and received approval from the United States Department of Education (USED) for administration of the Virginia Grade Level Alternative (VGLA) reading assessment for English Language Learners (ELLs) in grades 3-8 at proficiency levels 1 and 2. Participation in the VGLA is determined by the school-based ELL committee. The VGLA is an evidence-based alternative assessment for the Standards of Learning (SOL). Students who qualify to participate in the VGLA are required to demonstrate individual achievement of grade level content standards for the academic content area in which they are being assessed. Use of the VGLA for LEP students was permitted for AYP calculations based on tests administered in 2006-07 and for 2007-08.

LEP students who are at the lowest levels of English proficiency (levels 1 and 2) may participate in VGLA in reading ONLY for up to three consecutive years in lieu of taking the statewide SOL reading assessment. LEP students at level 3 and above must take the SOL reading assessment with or without accommodations. Virginia also has a "plain English" version of the mathematics tests for grades 3-8 and Algebra I—approved by ED through the peer review process of state assessment systems—where the language load of the test items has been reduced. LEP students whose language proficiency is at levels 1 and 2 may take the plain English math tests for up to 3 consecutive years.

**AYP Consequences, Reporting, and Appeals**

Summarized here are states' 2008 amendments aimed at modifying their educational accountability system designs with respect to the order of applying public school choice and supplemental educational services (SES) sanctions, state reporting timelines, and school or district appeals of AYP decisions.

**Public School Choice and Supplemental Educational Services**

Under NCLB, schools identified for improvement after missing AYP targets for two consecutive years must notify parents of that determination and, prior to the opening of school the following year, must offer students the option of public school choice, i.e., moving to another school. Schools missing AYP targets for a third consecutive year must offer supplemental educational services to the lowest-achieving low-income students (in addition to public school choice). Over the past several years, many states have sought to reverse the order of those sanctions, arguing that the latter are far more critical interventions for low-achieving students.

Until 2005, ED consistently denied states’ educational accountability workbook amendments to permit the reversal of public school choice and supplemental educational services sanctions for schools and districts.
identified for improvement under NCLB. In some cases, states had earlier enacted public school options and, like Utah, argued that most students were already attending schools of their parents’ choice. Others argued that it is more critical to provide supplemental educational services such as tutoring at the earliest possible time for low-achieving students. However, on June 4, 2008, ED suddenly announced that it would permit all states with fully (or nearing fully) approved assessment systems to reverse the order of public school choice and the provision of supplemental educational services (Mesecar, 2008, June 4). However, ED maintained the separate application process it established earlier outside of the review of accountability workbook revisions.

**Changing the Order of SES and Choice Sanctions**
The provision of some flexibility to reverse the order of these sanctions began in mid-2006, when Secretary Spellings (May 15, 2006) wrote to chief state school officers announcing a separate application process for states seeking approval to reverse the order of school choice and supplemental educational services (SES) for some or all of their schools identified for improvement. ED had approved a pilot project in 2005 permitting four school districts in Virginia to reverse the order of these sanctions. In late July and mid-August 2006, four states—Alaska (one district), Delaware (one district), Indiana (three districts), and North Carolina (seven districts)—received letters informing them that selected school districts in their states had been approved to offer SES in year one of school improvement. That announcement signaled an expansion of the pilot program launched in 2005 but did not provide details on which states applied and how the decisions were made. However, while permitting the state to continue its prior approval for four school districts to offer SES in year one, ED denied Virginia’s request for an expansion of the flexibility in 2006 because the state had not received approval for its final assessment system under NCLB.

For 2007-08, ED approved continuations from Alaska (expanded to two districts), Indiana (two continued districts and one new), and North Carolina (six continued and one new). Virginia was also approved to expand its initial pilot by adding five districts while two dropped out. The single district previously approved in Delaware dropped out of the pilot. Whether additional states sought to participate in the pilot program is unknown as are the criteria as to how ED made its final determinations in this area. In spite of the Secretary’s May 15, 2006, announcement, the pilot program had yet to be expanded either in terms of the number of states or school districts involved until the June 2008 announcement that the pilot would be opened to all states with approved (or almost approved) assessment systems.

In early August 2008, Secretary Spellings announced that Alabama, Arkansas, Tennessee, and Utah had been approved for 2008-09 to permit all schools districts in their respective states to reverse the order of choice and SES. At the same time, Alaska received approval to expand its participation to five school districts; North Carolina received approval to expand its participation to all school districts; and Virginia to 14 school districts. No mention was made regarding Indiana’s continued participation.

At the same time that the above states were first informed regarding their applications to reverse school choice and SES, at least for some school districts, the Secretary notified the Anchorage, Alaska, and Hillsborough County, Florida, school districts that they had been approved to join Boston and Chicago as SES providers, even though they had been identified for improvement under NCLB. ED extended approvals for these districts in 2007-08, and for Anchorage, Boston, Chicago, and Hillsborough County, Florida in 2008-09. Charlotte, North Carolina also received approval for 2008-09. How these school districts were originally selected to receive this flexibility, including information about any application and review process, remains unknown.
Last year, four States (Colorado, Connecticut, Pennsylvania, and Utah) submitted accountability workbook amendments requesting approval to reverse the order of public school choice and SES or to permit schools identified for improvement to offer both choice and SES in the first year. A fifth state—Virginia—sought to expand an earlier approval to reverse the order. All were denied. This year, three of those states have again submitted amendment requests.

In Pennsylvania, the state again sought approval to require schools missing AYP targets for two consecutive years to offer SES (instead of choice in year one) and both SES and public school choice for those missing targets for three consecutive years (previously denied by ED on July 2, 2007.). Pennsylvania’s primary argument remained the same—“The first year that a school falls into the Improvement Cycle, offering Supplemental Education[al] Services, such as tutoring, just makes sense. Offering Choice prior to tutoring does not address the issue of low performance at a school.” ED again denied the proposed amendment on August 18, 2008, “because it is inconsistent with section 1116 of the ESEA.”

Utah also re-submitted its earlier denied request (June 25, 2007) to permit schools to offer either choice or SES in the first year of identification for improvement. Utah has provided parents, by statute, with school choice for many years and because of this, the state notes, “children are, for the most part, already in the parent’s public school of choice.” However, on June 17, 2008, ED again denied the request indicating only that Utah could apply separately for participation in the Secretary’s pilot program allowing reversal of these sanctions; which the state did and received approval in August.

Like Pennsylvania and Utah, Virginia resubmitted the same request ED denied (July 16, 2007) to simply reverse the manner in which public school choice and SES requirements are applied to schools identified for improvement. Virginia noted that several school districts there have participated in ED’s pilot program and all report favorable results. As a result, the state “believes that all schools should have this option.” ED denied the request on June 24, 2008, while inviting the state, like Utah, to apply separately for participation in the pilot program.

**Targeting SES and Choice**

In 2006 and 2007, several states proposed targeting the SES and choice sanctions to specific student groups. However, this year, only two states submitted related amendment requests; both had submitted similar proposals in 2007 that ED denied.

Pennsylvania re-submitted its 2007 proposal related to applying differentiated consequences in cases where “a school/district fails to reach the threshold for only one of less than 10% of the measurable and appropriate AYP targets . . . .”. For the 2007 submittal, the state provided that schools and districts failing to reach the threshold for only one or less than 10% of the targets would not proceed in status but would be required to develop an improvement plan. For the 2008 submittal, Pennsylvania added a second component. In this case, schools and districts failing to meet between 10% and 40% of the AYP targets

---

44 In its denial, ED stated, “In the near future, the Department intends to open this pilot to all States who meet the eligibility requirements.” ED had not previously indicated an intention to make this policy shift nor did the letter provide any related details. Almost one year later, on June 4, 2008, the pilot was officially opened to all states. This followed the March 20, 2008, Secretary’s announcement of a limited pilot program related to “differentiated accountability” for up to 10 states (Spellings, 2008, March 20). This new pilot centers primarily on schools and districts identified for improvement and the application of consequences or sanctions in those schools and districts. (See Chapter 1 for a discussion of “differentiated accountability.”)
would proceed to the next accountability status but be exempt from public school choice, restructuring, and reconstitution. The state argued in both years that differentiated consequences afford more effective use of resources and not all schools and districts identified for improvement need to be subject to the same sanctions. Last year, Pennsylvania noted that of 24 school districts failing to make AYP in 2006, nine (37.5%) failed to reach the AYP thresholds for fewer than 10% of the measures; the same was true for 201 (36.7%) of the 548 schools identified for improvement. ED denied the request on August 18, 2008:

The Department cannot approve this proposal because it is inconsistent with section 1116 of the ESEA. Pennsylvania incorporated this proposed amendment into the proposal it submitted for the Department’s differentiated accountability pilot . . . peer reviewed on June 13-14, 2008. As was stated in the letter dated July 1, 2008, Pennsylvania’s proposal was not approved to be part of the differentiated accountability pilot at this time.

**Utah** re-submitted its third proposal to target SES and choice to the lowest-performing students, a request denied by ED in 2006 and 2007. ED stated in last year’s denial, “Section 1116(e) extends eligibility for SES to all students from low-income families. Only if a district needs to prioritize because it cannot meet all demand may it target these options to the lowest-achieving of these students.” On June 17, 2008, ED again denied the proposal and again cited section 1116(e) as the basis for the Department’s decision.

**School Districts Identified for Improvement as SES Providers**

As noted above, in 2004-05, ED directly allowed (bypassing the respective state education agencies) for the first time a few specific school districts identified for improvement to serve as supplemental educational services providers. However, that appears to be the only exception. In a letter to chief state school officers in the 10 Ed-Flex states (Colorado, Delaware, Kansas, Maryland, Massachusetts, North Carolina, Oregon, Pennsylvania, Texas, and Vermont), Assistant ED Secretary Kerri Briggs reaffirmed that those States do not have the authority to “waive the regulation in 34 CFR 200.47(b)(1)(iv)(B), which prohibits districts identified for improvement or corrective action from being providers of supplemental educational services (SES) under Part A of Title I.” She further stated:

I wanted to make it clear that the Ed-Flex legislation expressly prohibits States from waiving requirements related to sections 1111 or 1116 of the Elementary and Secondary Education Act (ESEA), as amended (i.e., requirements relating to State plans, standards, assessments, and local educational agency and district improvement). The legal basis for the regulation in 34 CFR 200.47(b)(1)(iv)(B) is section 1116 of the ESEA. Thus, States do not have the authority to waive this regulation.

I recognize that there has been some confusion over whether an Ed-Flex State may grant this type of waiver, and I hope that this letter will clarify what is allowable under the law. While the Department is not able to authorize such a State-wide waiver under Ed-Flex, I would suggest that you encourage

---

45 Ed-Flex is described at ED’s web site as, “a program that allows the Secretary of Education to delegate to states with strong accountability safeguards the authority to waive certain federal education requirements . . . that may, in particular instances, impede local efforts to reform and improve education. It is designed to help districts and schools carry out educational reforms and raise the achievement levels of all children by providing increased flexibility in the implementation of federal education programs in exchange for enhanced accountability for the performance of students.”
interested districts to apply for the Department's pilot program, which allows districts in need of improvement to provide supplemental educational services in certain circumstances. More information about this pilot can be found on our website: http://www.ed.gov/nclb/choice/help/sespilot-2006.html (Briggs, 2008, January 4).

**Reporting**

Although not submitting a specific educational accountability workbook amendment related to reporting, Montana raised an important issue facing many states with small populations—ensuring reporting confidentiality consistent with the Family Educational Rights and Privacy Act (FERPA) while striving to include all schools and districts in the state’s accountability system. According to Montana:

We are aware that NCLB Title I regulations (34 CFR Sec 200.7) prohibit reporting disaggregated achievement data on subgroups if that would disclose personally identifiable information about an individual student. However, reporting such data for all students (without subgroup disaggregations) in many of our schools would have the same result of disclosing personally identifiable information. Also, the same regulatory provision indicates that states must continue to make AYP determinations based on the performance of each subgroup. That seems inconsistent with the Title I statute, which defines AYP not to require disaggregations for subgroups if the results would reveal personally identifiable information about an individual student. In particular, we are unclear how we can make an AYP decision about a school that is based on information that would be personally identifiable and still provide school parents in the aggregate the statutory right to appeal a designation of their school as in improvement status. We also would appreciate any guidance as to whether there are ways to mask data that would meet privacy concerns but still adhere to the NCLB requirements.

After a surge in requests to delay reporting in 2006, caused in large part by states introducing additional assessments in grades 3-8 (the previous requirements were for one test in each of the 3-5 and 6-9 grade spans), the number of amendment requests in this area have tapered off. That year, nine states, sought a longer reporting timeline for AYP results primarily due to the additional time required to conduct standard-setting activities and revision of AYP algorithms to incorporate the expanded testing data. ED tended to approve these requests. Last year, only four states (Arizona, Kentucky, Michigan, and Oregon) submitted accountability system amendments related to delays in reporting. All were approved.

This year, five states submitted educational accountability workbook amendments related to reporting. New Jersey has redesigned four of its seven state assessments (grades 5-8) in language arts literacy and mathematics. The assessment window was also moved closer to the end of the school year. The state intends to re-set starting points, AMOs, and IGs. As a result, New Jersey indicated that the resetting process will delay AYP status notification based on 2008 assessment results. The state asked for approval to continue 2007 AYP status and sanctions for 2008-09 as it completes the necessary transition steps (approved July 16, 2008). Missouri clarified that it will meet all NCLB reporting requirements after 2008, including implementation of new end-of-course examinations at the high school level in 2008-09.

As noted earlier in this chapter, North Carolina administered new reading assessments in grades 3 through 8 in 2007-08 and requested approval to use only mathematics results for AYP decisions in 2008. Use of the new reading assessments would “necessitate a delayed release of reading results.” Using only mathematics results for AYP would “give ample time for the [reading] developmental scale to be determined, the new more rigorous achievement levels to be set . . . and the AMOs to be reset for reading.” The state did not provide documentation related to when the new reading scores would be ready to be reported. ED denied the request on June 6, 2008, and prescribed a transition plan similar to that described below for Texas. South Dakota clarified (approved July 15, 2008) that it will include
information about the professional qualifications of teachers there on the annual State Report Card. ED’s approval requires that “This information will also be included on the annual local educational agency report cards . . . [for those receiving] Title I funds."

**Texas** reported that it would conduct standard setting in August 2008 for its new alternate assessment based on modified student achievement standards. On May 19, ED approved a transitional plan for making AYP decisions until the preliminary standard setting results become available in October. ED stated:

> At the beginning of the 2008-09 school year, all schools and districts will have the same AYP status as during the 2007-08 school year. All schools and districts that cannot exit improvement status will implement the appropriate interventions prior to the start of the school year. If these schools and districts fail to make AYP based upon the 2007-08 assessment data, they will begin offering the next level of services. Schools and districts on the "watch list" (i.e., schools and districts that did not make AYP based upon 2006-07 assessment data) have been notified to plan for the possibility of offering services (i.e., public school choice) and must act immediately on the preliminary building assessment reports. Schools and districts previously identified as in need of improvement that make AYP for the first time will continue to offer the same level of services. Schools and districts previously identified as in need of improvement that make AYP for the second consecutive year based upon 2007-08 assessment data will no longer be identified for improvement and no action is required.

**Appeals to Identification for Improvement Determinations**

Early in NCLB’s implementation, ED did not seem to examine closely the manner in which states’ educational accountability workbooks described how they planned to address appeals from schools and districts with respect to identification for improvement. As noted in the 2003 paper (Erpenbach, Forte Fast, and Potts, pp. 27-28), “States were asked to describe, in Critical Element 9.2 of their accountability workbooks, their plans and procedures for schools and districts to appeal accountability decisions.” At that time, there was evidence of wide variation among states in their responses to this workbook element supporting the conclusion that this subject did not receive much attention from ED during the accountability workbook peer reviews earlier that year. This led to a caution in the 2003 paper that, “it is an area that could prove troublesome for states and school districts. The need for systematic, uniform, and objective processes/procedures for receiving and acting on evidence submitted by a school or district questioning an identification for improvement seems quite self-evident.”

By early 2007, it was clear that ED was beginning to take more interest in the matter of how states were handling appeals. As reported in last year’s paper (Erpenbach & Forte, 2007, p. 47):

> ED officials may be taking a closer look at how states are dealing with appeals of AYP determinations from schools and districts. For example, as a result of a Title I Monitoring Visit in April 2007, **Nevada** was directed to cease granting appeals from schools identified for improvement on the basis of making significant improvement from a lower to higher proficiency level. ED cited, in part, a section 1116(b)(2)(B) proviso that appeals must be based on “statistical or other substantive reasons.” ED’s decision did not elaborate on what “substantive reasons” might include nor does its treatment of this matter in the July 21, 2006, non-regulatory guidance on LEA and school improvement (see B-3, page 5 of the guidance).

Last year, two states submitted educational accountability workbook amendments concerning the handling of missing AYP target appeals. ED approved **Wyoming’s** clarification but denied **Arizona’s** request to establish as grounds for appeal evidence that a school’s failure to make AYP was due to the
inclusion of the scores of limited English proficient students in their first three years of schooling in the United States.

This year, only two states requested an amendment related to the handling of AYP decision appeals. **Rhode Island** sought approval for the state’s commissioner of education to be able to consider supplemental data as part of an appeals determination. The authority was to be “given tight boundaries” providing that, “Supplementary evidence appeals are only reviewed when a single target has been missed by a very small margin. The commissioner of education may review the appeal in the context of other performance indicators and make a judgment on whether AYP has been met.” On June 11, 2008, ED denied the request citing section 1116(b)(2)(B) which provides for appeals if the identification for improvement “is in error for statistical or other substantive reasons.” According to the department’s decision letter, “missing AYP by a small margin or missing AYP in only one subgroup is not a statistical error.” Again, ED did not provide examples of what might constitute “substantive reasons.”

**Utah** clarified the role of school districts in addressing appeals from schools (approved June 17, 2008). Under NCLB, the state has permitted appeals at the school and school district level, allowing 30 days for a school to appeal an AYP decision to the school district and a district 30 days to appeal an AYP decision to the state. Utah added:

- The school district determines whether a school is making adequate yearly progress.
- The procedures and guidelines for the appeal process are set by the Utah State Board of Education.

The state also wanted to add, “After the 30 day period, if the school district determines that an error would result in an incorrect AYP report for the following year, the district may appeal the report to ensure that subsequent designations are appropriate.” However, according to a Utah official, ED denied this although no mention is made of the matter in the department’s June 17, 2008, decision letter.

**The Next Chapter**

In Chapter 3, observations and conclusions are presented regarding states’ 2008 accountability workbook amendment requests and ED’s responses as posted at the department’s website. There are also comments on the few approvals that were somewhat surprising and likely next steps regarding states’ continuing efforts to strengthen their educational accountability systems. Whether this will continue to include the range and extent of accountability workbook amendments requests seen over the past two years as reauthorization of NCLB looms ever closer and a new administration takes over can only be a matter of speculation.
Chapter 3: Observations and Conclusions

In this chapter, states’ accountability amendments that were somewhat surprising to those who have followed the decision-making process over the past six years are described along with consistently denied requests. The observations and conclusions described here are again more informative this year in that ED has continued the practice begun last year of responding in writing to all 49 States and the District of Columbia submitting accountability workbook amendments, posting on its website both approved amendments and, in most cases, denied requests—sometimes with supporting rationale.

Unanticipated Approvals

This year, there were no truly unanticipated or “surprise” approvals. However, the approval of two distinct types of science testing at the high school level could be characterized as a minor approval surprise as was the fact that ED officials questioned at least one state on their participation rate determinations when formerly served LEP students were included in subgroup AYP measures.

Among this year’s notable educational accountability workbook amendment decisions were those such as:

- Approving the use of science assessments at the high school wherein not all students take the same assessment at the same grade. In these instances, ED approved the use of end-of-course (EOC) assessments in any of three or four science domains and then the state simply either rolled the scores together to create a sort of high school level annual science score or reported separate composite scores for each of the assessed domains. In this approach, not all students take the same EOC assessments, not all students take a science course every year, and students who take more science courses have more scores contributing annually.

- ED officials telling at least one state that they could use only achievement scores for formerly served LEP students when including them in that subgroup for AYP determinations (the state uses all three elements and has done so since 2004). The ED official cited non-regulatory guidance issued by the department in May 2007. That guidance includes at pages 11-12 a question and answer that taken literally would restrict a state to using solely achievement score data for formerly served LEP students. The non-regulatory guidance interpretation followed more than three years after Secretary Paige’s key policy letter during which time almost every state has incorporated the flexibility provided in its accountability workbook. The extent to which ED decides to pursue this issue with other states remains to be seen.

- One more puzzling than surprising with respect to amendments from New Mexico and Utah for use of a longitudinal safe harbor model similar to that approved for Colorado in 2007. In this case, ED granted Colorado a one-year approval last year that was extended this year for a longitudinal safe harbor model (in addition to use of the traditional model). The approval “allows any district, school, or student group to make adequate yearly progress. . . provided there is a 10 percent decrease in the percentage of non-proficient scores in the current year compared to the previous year, comparing individual student assessment results in the current and previous year.” The match rate for the group of students must be at least 95% or higher. ED cited the state’s data

---

46 As noted in the preceding chapter, ED’s April 2007 regulations related to modified academic achievement standards included summation comments and regulations that only the “scores” of formerly served SWDs could be included in SWDs student group AYP calculations. ED’s September 2006 regulations related to LEP students do not contact similar summation comments or regulations.
system that is able to track student achievement as a reason for the approval. This year, ED denied New Mexico’s request but approved part of Utah’s after initially rejecting the state’s amendment request. Last year, ED denied Louisiana’s request even though the state has a student information system that allows student-level data to be tracked over time.

- Approving what amounts to the use of four-year trend data in the determination of graduation rates in Alaska. In this case, Alaska will evaluate “the trend in graduation rates over the past four consecutive years, including the current year” whenever aggregating graduation rates for a school that still fails to meet the state’s minimum n. Oregon received a similar approval “when two years of data does not provide a sufficient number of tests.”
- Approving North Carolina’s somewhat of a “first” with respect to how a state makes AYP decisions in small schools. In this case, ED approved the state’s plan to permit school districts to provide a justification for the AYP designation instead of sending a team of educators to the district to make that determination. It should be noted that while ED stated in late 2002 that it would issue related non-regulatory guidance that does not yet appear to have happened.

**Consistently Denied Requests**

As in previous years, ED consistently rejected states’ requests to limit identification for improvement to schools and districts that missed AYP in the same subject and group for two consecutive years. New in 2006, and continued since, has been ED’s general rejection of proposals to increase minimum n’s, confidence intervals, or anything that appears to “tweak” AYP determinations such as the use of an index in lieu of percent proficient.

The Department also consistently denied targeting sanctions in schools identified for improvement to the student group(s) missing AYP targets (how consequences are implemented); any alternatives for the inclusion of ELLs in assessments or AYP; and the reversal of public school choice and the provisions of supplemental educational services unless the state pursued the reversal through a separate application process (although that did not guarantee approval).

**Next Steps**

In concluding the last two papers on accountability workbook amendments, it was opined that each of the coming school years would be a quiet one for accountability amendments as ED and many states had started shifting more of their focus to ESEA reauthorization discussions. As it played out, this forecast turned out to be wrong both times—states continue to spend time on fine-tuning their accountability systems—especially as they have brought on line additional assessments in grades three through eight, added science assessments, and otherwise made changes in their standards and assessments or added alternate assessments. Clearly, looming reauthorization and related issues have not dissuaded states from making changes to their accountability systems.

In 2008-09, states will have to have final approval of their science assessments in at least one grade in each of the 3-5, 6-9, and 10-12 grade spans. Many States will still be working toward approval of their assessments for students with significant cognitive disabilities. Many will also be developing modified achievement standards and assessments based on those standards. States will also be more likely to turn their attention to ESEA reauthorization issues, although many Washington “insiders” have already signaled their belief that actual reauthorization is unlikely before the end of 2009 or even into early 2010.
Among “trends” noted in ED’s decision letters this year were:

1. Continued cautions where states were perceived to have set low graduation rate targets; although not all states with seemingly low targets had cautions included in the decision letters they received from ED.

2. A caution regarding minimum n’s included in New Jersey’s (but not other states) decision letter that, “The Department intends to use the National Technical Advisory Council, recently established by Secretary Spellings, to review all states’ minimum group sizes and provide guidance to the Department in order to ensure valid and reliable accountability systems that, to the maximum extent possible, include all students.”

3. Signs the Department may be taking a closer look at exactly how states are making AYP calculations when formerly served, or exited, SWDs and ELLs are include in those student groups.

4. The department’s continued assertion that it will take a closer look at the use of indexing in AYP determinations. How this may affect the use of indexes and whether it results in related regulations or non-regulatory guidance remains unclear at this point.
References and Other Resource Material


Final Regulations, Title I—Improving the Academic Achievement of the Disadvantaged; Final Rule [regarding the achievement of students with the most significant cognitive disabilities], 68 Fed. Reg. 68698 (Dec. 9, 2003) (to be codified at 34 C.F.R. pt. 200).

Final Regulations, Title I—Improving the Academic Achievement of the Disadvantaged; Final Rule [regarding alternate assessments based on “modified” student academic achievement standards], 72 Fed. Reg. 17748 (April 9, 2007) (to be codified at 34 CFR Parts 200 and 300).


Hoff, D. J. (2008, April 2). State tests not all ok under law. Education Week, pp. 1, 24.


Johnson, H. L. (2006, March 7). Letter to chief state school officers regarding April 1, 2006, for the submittal of 2006 accountability workbook amendments and transmitting information related to the identification of LEAs for improvement and other information related to other aspects of state accountability system potentially impacted by the addition of grades required to be assessed in the 3-8 span beginning in 2005-06. Washington, DC: U. S. Department of Education.


NCLBonline (2008, June 24). NCLBonline posts compilation of public comments on Title I proposed rule. Washington, DC: Thompson Publishing Group, Inc.


Olson, L., & Hoff, D. J. (2006, December 13). Framing the debate. Education Week, pp. 22-30.


Rooney, P. (personal communication, November 16, 2007)


APPENDIX A – Acronyms Used In This Paper

AMOs          Annual Measurable Objectives
AMAOs         Annual Measurable Achievement Objectives
ASR SCASS     Accountability Systems and Reporting State Collaborative on Assessment and Student Standards
AYP           Adequate Yearly Progress
CIs           Confidence Intervals
CCSSO         Council of Chief State School Officers
ED            United States Department of Education
ELL           English Language Learner
ELP           English Language Proficiency
EOC           End-of-Course Test
ESEA          Elementary and Secondary Education Act of 1965
ESL           English as a Second Language
FAY           Full Academic Year
IASA          Improving America’s Schools Act of 1994
IDEA          Individuals with Disabilities Education Act
IEP           Individualized Education Program
IGs           Intermediate Goals
LEAs          Local Educational Agencies (also referred to as school districts)
LEP           Limited English Proficient
NAEP          National Assessment of Educational Progress
NCLB          No Child Left Behind Act of 2001
NGC           National Governors Association
NPRM          Notice of Proposed Rulemaking
OAlIs         Other Academic Indicators
SEA           State Educational Agencies (also referred to as State Departments of Education)
SES           Supplemental Educational Services
SWDs          Students with Disabilities
APPENDIX B—Timeline of Significant Events Related to State Accountability Plans Under NCLB

**2002**
- January—The No Child Left Behind Act of 2001 becomes law.
- July—ED issues *Standards and Assessment* Regulations.
- December (early)—ED issues *Educational Accountability* Regulations.
- December (late)—ED releases Accountability Workbook Templates to states.

**2003**
- January 31—Draft Accountability Workbooks due to ED.
- January – April—Peer Reviews of State Accountability Workbooks conducted in states.
- June 27—Secretary’s Letter regarding inclusion of SWDs in AYP.
- June (late)—All State Accountability Plans tentatively approved by ED (several with compliance agreements).
- July—CCSSO releases Year One paper on State Accountability Plan requests and ED approval decisions.
- August—States begin submitting Accountability Plan amendment requests to ED; a process that would become on-going.
- December 9—ED issues *Achievement of Students with the Most Significant Cognitive Disabilities* Regulations.
- December 11—Secretary’s Letter regarding additional flexibility for SWDs in state assessment and accountability systems.
- December 18—ED releases Q & A regarding alternate achievement standards.

**2004**
- February 5—Assistant Secretary’s Letter setting April 1 deadline for accountability amendment requests applying to 2003-04 AYP decisions.
- February 20—Secretary’s Letter announcing new flexibility related to ELLs.
- March 2—Secretary’s Letter announcing further flexibility related to AYP for SWDs.
- March 29—Secretary’s Letter regarding calculation of participation rate.
- April 7—ED issues *Standards and Assessments Peer Review Guidance* for the NCLB Peer Reviews.
- May 21—Assistant Secretary’s Letter elaborates on flexibility for participation rate.
- May 21—Second Secretary Letter’s regarding calculation of participation rate.
- October—CCSSO releases Year Two paper (2003-04) on State Accountability Plan Amendments.

**2005**
- January 27—Assistant Secretary’s Letter setting April 1 Deadline for 2004-05 Accountability Plan Amendments.
- February 16-19—ED conducts Standards and Assessments Peer Reviewer Training and first round of Standards and Assessments Peer Reviews under NCLB.
- April 7—Secretary Spellings announces *Raising Achievement: A New Path for No Child Left Behind* initiative.
- May 10—Secretary sets June 1 deadline for state accountability amendments and issues information related to serving SWDs with “Persistent Academic Difficulties.”
June 13—ED issues updated non-regulatory guidance concerning the provision of supplemental educational services.

August—ED releases non-regulatory guidance concerning alternate achievement standards for students with the most significant cognitive disabilities.

September 29—Secretary’s Letter concerning students displaced by Hurricanes Katrina and Rita.

October—CCSSO releases Year Three paper (2004-05) on State Accountability Plan Amendments.

December—NPRM released regarding modified achievement standards and assessments for some SWDs.

2006

January 25—ED issues guidance for NCLB Growth Model proposals and sets February 17 application deadline. Eight states approved for peer review in April and two (Tennessee and North Carolina) approved to implement model in May.

March 7—Assistant Secretary’s Letter setting April 1 deadline for new accountability amendment requests and related transition information.

May 15—Secretary’s Letter announcing separate application process for participation in extension of pilot program permitting schools to reverse order of choice and SES and some LEAs identified for improvement to serve as SES providers.

June—all 49 yet-to-be-approved states (50 states, Puerto Rico, and the District of Columbia) notified of assessment system peer review results.

July 21—ED releases revised non-regulatory guidance related to LEA and school improvement.

September 13—ED issues new regulations related to the achievement of LEP students.

September 19—Assistant Secretary’s Letter updating chief state school officers on ED’s priorities.

November—CCSSO releases Year Four paper (2005-06) on State Accountability Plan Amendments.

November—ED announces that Arkansas, Delaware, and Florida are approved for growth models following submittal of revised plans in September.

2007

February 7—Secretary’s Letter updating chief state school officers on matters related to the implementation of Title I assessment and accountability requirements.

March—second round of peer reviews for growth model proposals conducted. Four states—Alaska, Arizona, Iowa, and Ohio—eventually approved.

April—CCSSO releases paper on alternate assessments for SWDs with significant cognitive disabilities.

April 9—ED issues regulations related to modified achievement standards and other matters as well as draft non-regulatory guidance related to the modified achievement standards.

May—ED releases non-regulatory guidance for the assessment and accountability of LEP students.

May—ED releases An Update on State Standards and Assessment Systems.

June 27—Secretary conducts National Summit on Supplemental Educational[al] Services.

June 27—Assistant Secretary’s Letter to chief state school officers announcing department’s intent to review the use of performance indexes in AYP determinations.

July 25—ED conducts invitational meeting for states without approved alternate assessments for students with significant cognitive disabilities.

July 26-27—ED hosts meeting for interested states to discuss development of standards and assessments based on modified achievement standards.
• September 5—Secretary’s Letter to chief state school officers conveying information about the department’s review of state assessment systems. Includes a listing of the peer reviewers.
• November 17—Assistant Secretary’s Letter to chief state school officers announcing process and timelines for amending accountability workbooks in 2008.
• December—CCSSO releases Year Five paper (2006-07) on State Accountability Plan Amendments.
• December 7—ED announces that all states with approved assessment systems may now use growth models in their AYP designs.
• December 21—ED releases revised Standards and Assessments Peer Review Guidance to include modified academic achievement standards.

2008
• February 28—Assistant Secretary’s Letter to chief state school officers announcing science standards and assessments peer reviews.
• March 10—Secretary announces differentiated accountability model pilot program aimed at how sanctions may be applied when schools and districts are identified for improvement under Section 1116.
• April 23—Secretary announces Notice of Proposed Rulemaking related to nine areas of NCLB including graduation rates, minimum group size, and same subject-same student group identification for improvement (prohibiting same).
• May 2—ED issues Notice of Proposed Interpretations intended to ensure uniformity across the states in matters related to the assessment of English language acquisition and when LEP students may exit special services programs.
• June—ED announces that Michigan and Missouri are approved to use growth models.
• June 4—ED announces that pilot program permitting school districts to reverse the order of public school choice and the provision of supplemental educational services will be open to all states with approved statewide assessment systems.
• July—Secretary announces that Florida, Georgia, Illinois, Indiana, Maryland, and Ohio are approved to participate in the differentiated accountability model pilot program.
• August 4—Secretary Spellings announces that seven states—Alabama, Alaska, Arkansas, North Carolina, Tennessee, Utah, and Virginia—are the first states approved under the pilot program permitting school districts to reverse the order of public school choice and supplemental educational services in schools sanctioned under NCLB.
• September 16—first meeting of ED’s National Technical Advisory Council
• October—CCSSO release Year Six paper (2007-08) on State Accountability Plan Amendments
## APPENDIX C – State Educational Accountability Workbook Decisions by the U. S. Department of Education that Changed Between 2003 and 2008

<table>
<thead>
<tr>
<th>Topic</th>
<th>Original Decision</th>
<th>Final Decision(^{47}) and Year Changed</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Accountability systems—dual</td>
<td>Permitted only when a school could not achieve a high performance level on state system if it was identified for improvement under NCLB.</td>
<td>States can have dual systems that recognize school’s performance regardless of their AYP outcomes (2003).</td>
<td>ED approved used of the term “Provisional AYP” by a state as a category for schools ranking at the top of its accountability system but not making AYP under NCLB.</td>
</tr>
<tr>
<td>2. Alternate assessments for SWDs</td>
<td>Permitted but must be based on the state’s academic content and student academic achievement standards.</td>
<td>Alternate assessments based on alternate achievement standards permitted, but not more than 1% of total student population in the assessed grades can be counted as proficient in relation to the alternate standards (2004). Alternate assessments based on modified achievement standards permitted, but not more than 2% of total school population (in addition to 1% above) can be counted as proficient in relation to the alternate standards (2005).</td>
<td>Modified in regulations adopted on December 9, 2003 (Federal Register). See Paige (2004, March 2) policy letter on SWDs. Announced by Secretary Spellings in April 2005 with guidelines issued May 10, 2005. Codified in regulations promulgated April 2007.</td>
</tr>
</tbody>
</table>

\(^{47}\) A state that would like to amend its plan based on these decisions would be required to submit a request for amendment to ED; that is, states cannot simply amend their plans without approval even if other states have been approved to make the same amendment. The decisions should not be considered as "precedent setting." That is, approval for one state does not necessarily mean that another state will receive approval for the same amendment.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Original Decision</th>
<th>Final Decision and Year Changed</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. AYP: Decisions for school districts</td>
<td>Districts required to “roll up” data across schools for AYP determinations.</td>
<td>Districts may examine AYP outcomes by grade spans (elementary, middle, and high school) and then be identified for improvement only when the same content area target is missed in all spans for two consecutive years (2004). ED clarified in 2006 that missing OAI must be factored in as well.</td>
<td>See Johnson (2006, March 7) letter to chief state school officers.</td>
</tr>
<tr>
<td>4. AYP: Identification for improvement</td>
<td>School or district missing any AYP target two consecutive years is identified for improvement.</td>
<td>School or district must miss any target in the same subject or the same other academic indicator in two consecutive years to be identified for improvement (2004).</td>
<td>ED has consistently denied states’ requests to base identification on the same target two consecutive years. Proposed NPRM issued April 23, 2008, would codify the prohibition.</td>
</tr>
<tr>
<td>5. AYP: Rounding up decisions</td>
<td>Unclear.</td>
<td>Some states permitted to round up when calculating AYP elements (2004).</td>
<td></td>
</tr>
<tr>
<td>6. AYP: Use of first test score when students have multiple opportunities to take an assessment</td>
<td>First test score counts for AYP determinations.</td>
<td>States may “bank” results when students afforded multiple opportunities to test provided an “official” point has been designated at which they are expected to have attained the tested standards (2003).</td>
<td>ED stated in comments with Regulations adopted on December 9, 2003, (Federal Register) that states have more flexibility toward this end than originally understood. In April 2007, ED rescinded the regulatory requirement pertaining to the use of the first test score.</td>
</tr>
<tr>
<td>Topic</td>
<td>Original Decision</td>
<td>Final Decision and Year Changed</td>
<td>Comments</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------</td>
<td>--------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Delaware</td>
<td>approved to recalculate AYP following summer school assessment &quot;retakes&quot; (2006); Virginia to include retesting results for high school end-of-course assessments (2006); and Michigan in 2007 to include re-test scores due to the &quot;high stakes&quot; nature of its high school assessments.</td>
<td>ED clarified in a 2007 decision letter to Pennsylvania that a state may not include retesting results for high school assessments beyond the grade level for which the assessment is aligned to the state's academic content standards.</td>
<td></td>
</tr>
<tr>
<td>Virginia</td>
<td></td>
<td>Three additional states—Georgia, Maryland, and Tennessee approved in 2008.</td>
<td></td>
</tr>
<tr>
<td>7. ELLs: Including scores for “exited” ELLs in the ELL student group AYP decisions</td>
<td>Not permitted.</td>
<td>Permitted for up to two years after student no longer meets the definition of LEP under section 9101(25) consistent with the Secretary’s policy letter (2004).</td>
<td>See Paige (2004, February 20) policy letter on ELLs. States using Title IX definition of LEP could include such students for longer than two years.</td>
</tr>
<tr>
<td>8. ELLs: Use of achievement scores in AYP calculations</td>
<td>Required.</td>
<td>Recent arrivals in the U. S. may be exempted from one administration of a state’s reading or language arts assessments. As long as these students take an English language proficiency assessment, they may be counted as a participant. Recent arrivals cannot be exempted from a state’s mathematics and science assessments but mathematics scores may be exempted from one cycle of AYP determinations (science results are not required to be included in AYP measures).</td>
<td>See Paige (2004, February 20) policy letter on ELLs. See also September 2006 regulations.</td>
</tr>
<tr>
<td>Topic</td>
<td>Original Decision</td>
<td>Final Decision and Year Changed</td>
<td>Comments</td>
</tr>
<tr>
<td>-------</td>
<td>------------------</td>
<td>--------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>9. Full academic year extending beyond one calendar year for SWDs</td>
<td>Not permitted.</td>
<td><strong>Iowa</strong> approved (2004) to extend Full Academic Year for SWDs beyond one year when IEP indicates that longer period is needed to collect achievement information.</td>
<td>Proposed NPRM issued April 23, 2008, would modify many of these considerations.</td>
</tr>
<tr>
<td>10. Graduation rate covering more than four years</td>
<td>Students taking more than four years to graduate could not be counted as graduates in calculating graduation rates.</td>
<td>SWDs who take more than four years to graduate, consistent with their IEPs, may be counted as graduates (2004).&lt;br&gt;ELLs who take up to six years to graduate, consistent with a local plan, may be counted as graduates (2005).&lt;br&gt;<strong>Tennessee</strong> approved (2007) to define standard number of years as five plus one summer session for early college/middle college high schools. <strong>Michigan</strong> received a similar approval that year.</td>
<td></td>
</tr>
<tr>
<td>11. Growth models as accountability measures</td>
<td>Not permitted.</td>
<td>Pilot growth models (up to 10) permitted subject to peer review beginning in 2006. December 2007—Secretary opens the pilot to all states with approved assessment systems.</td>
<td>Two states approved to implement models in 2006-07; seven additional states approved in late 2006 and early 2007; two additional states in June 2008.</td>
</tr>
<tr>
<td>12. Minimum n’s for school districts—larger</td>
<td>Not permitted initially. In 2006, policy changed back to non-approval.</td>
<td>Models approved permitting the use of a minimum n that increases proportionally as the size of the district increases—subject to caps (2003).&lt;br&gt;<strong>Colorado</strong> approved (2008) to use 30 or 1% when enrollment exceeds 3,000.</td>
<td>Not permitted in 2006.</td>
</tr>
<tr>
<td>Topic</td>
<td>Original Decision</td>
<td>Final Decision and Year Changed</td>
<td>Comments</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>In 2006, policy changed back to non-approval.</td>
<td>Models permitted with even larger minimums that increase proportionally as enrollments across</td>
<td>April 2007 regulations effectively retracted previous approvals to use larger minimum n’s for student groups.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tested grades increases—subject to caps (2005).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Not permitted in 2006.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Students with invalid test scores must be counted as non-participants in calculating this rate</td>
<td></td>
</tr>
<tr>
<td>15. Percent proficient determinations</td>
<td>Use number of students enrolled full academic year for the denominator.</td>
<td>Use number of students enrolled for a full academic year and tested for the denominator (2004).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>In June 2007, ED signaled its intent to review the use of indexing for AYP decisions.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>In June 2008, ED denied two states’ requests for growth models because the states used indexing</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>in their AYP decision rules.</td>
<td></td>
</tr>
<tr>
<td>Topic</td>
<td>Original Decision</td>
<td>Final Decision and Year Changed</td>
<td>Comments</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>---------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>17. Public school choice and supplemental educational services: Reverse order of application when a school is identified for improvement</td>
<td>Not permitted</td>
<td>Permitted on a pilot basis for some LEAs (2005).</td>
<td>Secretary announced “differentiated accountability model” pilot program in March 2008. Program aimed at a differentiated approach to the application of NCLB sanctions for schools and districts identified for improvement.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Opened to all states with approved assessment systems in June 2008.</td>
<td></td>
</tr>
<tr>
<td>18. Safe harbor: Averaging for up to three years for percent proficient determinations</td>
<td>Not permitted</td>
<td>Permitted (2005). Percent not proficient would have to decline by 10% over one year, 19% over two years, and 27% over three years.</td>
<td></td>
</tr>
<tr>
<td>20. Safe harbor: Reviews for small schools in percent proficient determinations</td>
<td>Unclear.</td>
<td>A state may conduct a review of current year and prior year performance even when its minimum n requirements are not met (2004).</td>
<td>States have been allowed to apply confidence intervals (CIs) to AYP calculations in general. But when requests for CIs have been specifically around “count” items, they have been denied.</td>
</tr>
<tr>
<td>21. Statistical tests—for participation and graduation rates</td>
<td>Not permitted for “count” elements.</td>
<td>Permitted (2004).</td>
<td>States have been allowed to apply confidence intervals (CIs) to AYP calculations in general. But when requests for CIs have been specifically around “count” items, they have been denied.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>In 2006, policy changed to non-approval for requests to increase CIs (e.g., 95% to 99%).</td>
<td>States have been allowed to apply confidence intervals (CIs) to AYP calculations in general. But when requests for CIs have been specifically around “count” items, they have been denied.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>In 2008, ED denied Tennessee’s request to use a CI in graduation rate calculations.</td>
<td>States have been allowed to apply confidence intervals (CIs) to AYP calculations in general. But when requests for CIs have been specifically around “count” items, they have been denied.</td>
</tr>
<tr>
<td>22. Statistical tests—in safe harbor reviews</td>
<td>Not permitted</td>
<td>Use of a 75% confidence interval allowed (2004).</td>
<td>States have been allowed to apply confidence intervals (CIs) to AYP calculations in general. But when requests for CIs have been specifically around “count” items, they have been denied.</td>
</tr>
<tr>
<td>Topic</td>
<td>Original Decision</td>
<td>Final Decision and Year Changed</td>
<td>Comments</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------</td>
<td>------------------------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>23. Supplemental educational services: Permit LEAs identified for improvement to be providers</td>
<td>Not permitted.</td>
<td>Permitted on a pilot basis (2005).</td>
<td></td>
</tr>
<tr>
<td>24. SWDs: Including scores for “exited” SWDs in student group AYP decisions</td>
<td>Not permitted with two exceptions—those for Georgia (2003) and South Carolina (2004).</td>
<td>In regulations promulgated in April 2007, states may now include exited SWDs in the student group for up to two years after receiving services.</td>
<td>The two-year limit may contradict IDEA provisions related to the provision of monitoring services to formerly served SWDs.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>See also Illinois’ 2006 decision letter from ED for further guidance.</td>
</tr>
<tr>
<td>27. Testing—out-of-level</td>
<td>Not permitted.</td>
<td>Permitted as an alternate assessment aligned to alternate achievement standards and subject to 1% cap for SWDs reported at the proficient or higher level (2004).</td>
<td>Modified in Regulations adopted on December 9, 2003 (Federal Register). See Paige (2004, March 2) policy letter on SWDs.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Prohibited as an alternate assessment based on “modified” achievement standards in regulations issued April 2007.</td>
<td></td>
</tr>
<tr>
<td>Topic</td>
<td>Original Decision</td>
<td>Final Decision and Year Changed</td>
<td>Comments</td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>28. Uniform averaging</td>
<td>Uniform averaging required.</td>
<td>States permitted to apply “non-uniform” averaging such that schools and districts may use the scores from only current school year or average across up to three years, whichever benefits the school or district more (2003).</td>
<td>Alaska approved to use 4-year averaging (2008).</td>
</tr>
</tbody>
</table>