

Revisiting Statewide Educational Accountability Under NCLB

A Summary of State Requests in 2003-2004
for Amendments to State Accountability Plans

*A paper commissioned by the CCSSO Accountability
Systems and Reporting State Collaborative*

Ellen Forte Fast
William J. Erpenbach



The Council of Chief State School Officers
One Massachusetts Avenue, NW, Washington, DC

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.

DIVISION OF STATE SERVICES AND TECHNICAL ASSISTANCE

The Division of State Services and Technical Assistance supports state education agencies in developing standards-based systems that enable all children to succeed. Initiatives of the division support improved methods for collecting, analyzing and using information for decision-making; development of assessment resources; creation of high-quality professional preparation and development programs; emphasis on instruction suited for diverse learners; and the removal of barriers to academic success. The division combines existing activities in the former Resource Center on Educational Equity, State Education Assessment Center, and State Leadership Center.

STATE COLLABORATIVE ON ASSESSMENT AND STUDENT STANDARDS

The State Collaborative on Assessment and Student Standards (SCASS) Project was created in 1991 to encourage and assist states in working collaboratively on assessment design and development for a variety of topics and subject areas. The Division of State Services and Technical Assistance of the Council of Chief State School Officers is the organizer, facilitator, and administrator of the projects. SCASS projects accomplish a wide variety of tasks identified by each of the groups including examining the needs and issues surrounding the area(s) of focus, determining the products and goals of the project, developing assessment materials and professional development materials on assessment, summarizing current research, analyzing best practice, examining technical issues, and/or providing guidance on federal legislation.

Funding support for this paper was entirely from member states of the Accountability Systems and Reporting State Collaborative (ASR-SCASS). For information about ASR-SCASS and for 50-state Accountability Profiles, see the CCSSO website, http://www.ccsso.org/projects/Accountability_Systems/.

2004

COUNCIL OF CHIEF STATE SCHOOL OFFICERS

David P. Driscoll (Massachusetts), President

Douglas D. Christensen (Nebraska), Vice President

G. Thomas Houlihan, Executive Director

Julia Lara, Deputy Executive Director, Division of State Services and Technical Assistance

Rolf K. Blank, Director of Education Indicators

Copies of this report may be ordered for \$10 per copy from:

Council of Chief State School Officers

One Massachusetts Avenue, Suite 700

Washington, DC 20001

Phone: 202-336-7016, Fax: 202-408-8072

www.ccsso.org

Revisiting Statewide Educational Accountability Under NCLB

A Summary of State Requests in 2003-2004
for Amendments to State Accountability Plans

Ellen Forte Fast
William J. Erpenbach

October 2004

*Accountability Systems and Reporting State Collaborative
on Assessment and Student Standards*
The Council of Chief State School Officers, Washington, DC

ISBN: 1-884037-93-3

Copyright © 2004 by the Council of Chief State School Officers. All rights reserved.

Table of Contents

Executive Summary	iii
Summary of Accountability Amendment Requests from States To Date	iv
Requests Consistently Not Approved	v
Section I: Introduction and Background	1
Section II: Key Issues in States' Amendments to their	
NCLB Accountability Plans in 2003-04	3
Standards and Assessments	3
Changes to the Test or Grade Level Used for AYP	3
Banking of Test Results	4
Out-of-Level Tests or Other Alternate Assessments	5
Adjustment of Scores Designating Proficiency ("Cut Scores")	6
Adequate Yearly Progress (AYP) Model	6
AYP Application	7
AYP Indicators	11
Annual Measurable Objectives, Intermediate Goals, and Trajectories	15
Strategies Intended to Enhance Reliability	16
Inclusion of All Students in Accountability	20
AYP Consequences and Reporting	22
Timing and Degree of Consequences	22
Rewards and Recognition	23
Changes to Program or Demographic Variables	23
Reporting Timeline	24
Section III: Conclusions and Observations	25
Requests Consistently Not Approved	25
Unanticipated Approvals	26
Requests that Await Public Response	26
Next Steps	27
References	29
Appendix A	31
State Accountability Workbook Decisions That Changed Between 2002 and 2004	31
Appendix B	35
Ten Principles for State Accountability Systems	35

Executive Summary

A major requirement for states under the *No Child Left Behind Act of 2001* (NCLB) was to plan and implement a statewide accountability system that met highly specific criteria. Each state submitted its initial draft accountability plan, called an “educational accountability workbook,” to the U.S. Department of Education (ED) in January 2003. Following a peer review process and responses from ED, all states re-submitted their accountability plans with required revisions and received approval for these plans by June 2003. Shortly thereafter, the Council of Chief State School Officers (CCSSO) prepared and released an initial report that examined common issues and themes in states’ accountability workbooks and ED’s ensuing approval decisions (Erpenbach, Forte Fast, & Potts, 2003).

The purpose of the present paper, is to summarize the predominant issues presented by states in their requests to amend their state accountability plans to date. By late summer 2003, some states had already begun submitting amendments to their accountability workbooks; the pace of submissions accelerated through the spring and summer of 2004. A number of factors may have contributed to this acceleration, including the high stakes consequences of the NCLB accountability requirements, the additional federal flexibility allowed for several provisions of the requirements, delays in the provision of regulations and guidance from ED, exchange of information among states on elements of their plans, and the initial release of 2002-03 AYP results. With 47 states submitting amendment requests by the April 1, 2004 deadline set by ED, the Accountability Systems and Reporting State Collaborative (ASR-SCASS) and CCSSO recognized the need for a paper to summarize the core issues states were trying to address with the amendments.

Information supporting this summary and analysis was provided to CCSSO by state education officials who voluntarily forwarded their amendment requests and approval letters from ED. Some information was also obtained from websites, with follow-up calls to staff of state departments of education. To protect states’ privacy, states are only named in this summary when the state request was described in a public official response letter from ED; other descriptions of state requests are kept “state blind.”

This paper is intended to summarize areas of the NCLB legislation’s accountability requirements where states have sought substantive changes to their accountability plans or have sought targeted technical assistance. Complete information on state accountability systems is reported by states on the CCSSO web site through 50-state accountability profiles (<http://accountability.ccsso.org/>). The website includes up-to-date information on state accountability systems and incorporates state accountability plan amendments for 2003-04.

To facilitate interpretation, this paper generally follows the same structure as its predecessor (Erpenbach, Forte Fast, & Potts, 2003). The content of this report is organized as follows:

- Section I presents an introduction and develops the background related to this paper.
- Section II describes the central issues addressed in the amendments proposed by states.
- Section III offers a set of conclusions reached as a result of our review of states’ accountability workbook amendments and ED’s responses. Also identified are “unanticipated approvals” and amendment requests awaiting ED decisions.
- Appendix A provides a summary of ED’s State Accountability Workbook Decisions that Changed Between 2002 and 2004.
- Appendix B explains the 10 principles states were asked by ED to describe in their accountability plans.

Neither CCSSO nor the authors necessarily endorse any of the educational accountability strategies embedded in the amendments described in this paper. The purpose of this paper is to describe amendment proposals, not to judge any state’s intent in making specific proposals or whether these 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.

Summary of Accountability Amendment Requests from States To Date

This report organizes into four main categories the requests for amendments to accountability systems and plans submitted by states to ED by the April 1, 2004, submission deadline. The categories were created for the purpose of summary and discussion in this paper. The outcomes of these requests are discussed in the body of the paper.

- 1) **Standards and Assessments:** One-fourth of states submitted amendment requests related to their standards and assessments, including the following:
 - Changing the test or grade level that yields scores used in AYP analyses;
 - "Banking" of test results for use in AYP determinations;
 - Using of out-of-level tests or other alternate assessments;
 - Adjusting assessment scores for defining "proficient."
- 2) **AYP Model:** Almost all states that submitted amendments requested at least one change to their AYP model. These requests can be categorized as follows:
 - AYP application rules including requiring two years of missing targets for identifying schools and districts, retroactive application, and defining Title I students;
 - AYP indicators including the use of rounding rules and defining percent proficient on required assessments, student participation rate in assessments, and other academic indicators such as the graduation rate;
 - State targets and schedules for AYP including establishment of annual measurable objectives and intermediate goals;
 - Strategies to enhance reliability of AYP results, including establishing minimum "n" sizes, using confidence intervals, and applying safe harbor reviews.
- 3) **Inclusion of All Students in Accountability:** Almost all states requested amendments to take advantage of the flexibility issued by ED regarding the following:
 - Limited English Proficient (LEP) students' assessment and inclusion in accountability systems;
 - Students with disabilities' inclusion in alternate assessments and accountability determinations.
- 4) **AYP Consequences and Reporting:** A small number of states requested amendments in this area, such as the following:
 - Timing of consequences for schools;
 - Implementing state rewards and recognition;
 - Organizing for reporting, including integration with state reports.

Requests Consistently Not Approved

To date, ED has been consistent in not approving several categories of requested changes from states. ED responses regarding the items on the following list have been consistent. These items can be considered “non-negotiable” types of requests, as they have been consistently denied:

- Using out-of-level testing results in AYP calculations, except when such practice is consistent with the 1 percent cap requirements;
- Limiting identification of schools as in need of improvement to those schools that miss AYP for two consecutive years in the same content area and student subgroup;
- Applying retroactively any 2003-04 amendments to years prior to the 2003-04 school year;
- Exempting schools from the obligation to test students with significant medical-emergencies via a blanket policy (decisions must be made on a case-by-case basis);
- Switching the order of the NCLB-prescribed sanctions of school choice and supplemental educational services—school choice must be offered in the first year of identification. Further, school choice cannot be limited to “non-proficient” students.

In the following sections of the paper, the authors set the context for the discussion of state amendments to their accountability systems, describe specifics about the approach and rationale for selected examples of state requests that were approved, and provide the rationale given by ED for requests not approved (where that rationale is known).

Section I: Introduction and Background

Under the No Child Left Behind Act of 2001 (NCLB), all U.S. states, the District of Columbia, and Puerto Rico were required to submit State Educational Accountability Workbooks (workbooks) to the U.S. Department of Education (ED) by January 31, 2003¹. Each workbook was organized using 10 categories, as established by ED. (Appendix C includes a list of these 10 principles.) For example, all states were required to present a plan for a statewide accountability system that included all public schools (K-12) and included all students enrolled in a school. For many states that had already established statewide accountability systems based on state policies, the transition to the specific NCLB requirements involved adaptation of existing systems and often change in guiding assumptions and methodology. For almost all states the NCLB requirements involved expansion and development of assessments, data system capacity, and methods for reporting and dissemination of results.

The initial workbooks comprising state plans presented by states were reviewed by peer teams and by ED staff in spring 2003. The general process following these reviews involved a series of negotiations and interactions conducted between state leaders and federal officials. All 52 revised workbooks were tentatively approved by June 2003. In most cases, ED required states to make some changes to their workbooks prior to granting approval. In all cases, approval remained contingent on implementation of the plans as described in the workbooks. (See ED language included in direct correspondence to states regarding their accountability plans, available at <http://www.ed.gov/admins/lead/account/letters/index.html>.)

Even before gaining this initial approval for their plans, states had to begin implementing many aspects of their plans in order to meet the requirement for making and publicizing their adequate yearly progress (AYP) decisions prior to the start of the 2003-04 school year. For example, states had to communicate their accountability plans to their stakeholders; adopt protocols for preparing their assessment, graduation, and other academic indicator data for AYP analyses; develop reporting mechanisms that met the NCLB requirements; develop plans to support students taking advantage of mandated school choice options; and develop processes and staffing for handling anticipated appeals to AYP determinations. Most states invested substantial staff time to these tasks during the spring and summer of 2003. After this period, many shifted a portion of their investments to the study of that first NCLB AYP cycle: examining what worked and what required refinement. This led most to think about modifications to their accountability systems; a process that was further stimulated by the issuance of regulatory changes in winter 2003-04.

States that make “significant changes” to their plans under Title I—including their assessment and accountability systems—may request federal review of those changes at any time (sec. 1111(f)(2)). Toward this end, some states began submitting proposed amendments as early as summer 2003. Anticipating that many states intended to amend their educational accountability systems and desiring to streamline the process, the newly appointed Assistant Secretary for Elementary and Secondary Education, Raymond Simon, sent a letter on February 5, 2004, to chief state school officers outlining procedures for how states could request federal approval for amendments to their accountability plans. This letter set a deadline of April 1, 2004, for submission of amendments that states intended to apply to their 2003-04 data in order to make 2004-05 AYP decisions. Many states took advantage of this opportunity and Simon pledged that ED would “make every effort” to respond to any requests within 30 days². These requests for accountability plan amendments and the ensuing federal responses³ are analyzed in this paper.

¹ For more complete information on the initial submission and review process, see the recent CCSSO report (Erpenbach, Forte Fast, & Potts, 2003), available from CCSSO publications at <http://www.ccsso.org/publications/details.cfm?PublicationID=215>.

² According to an *Education Week* article, “. . . while department officials have been in constant contact with states, by late last week they had given final, written approval to only 20 of the approximately 45 that submitted requests by April 1” (Olson, 2004, July 14). By August 2004, ED had responded to many more of the plan amendment requests.

³ Response letters to the states are available at <http://www.ed.gov/policy/elsec/guid/stateletters/index.html>.

The present paper summarizes states' requests for amendments to their educational accountability plans and does not represent a comprehensive statement of the current status of these plans, nationwide. Many states requested changes that were already part of other states' approved plans, so the amendment requests do not necessarily reflect the prevalence of any specific strategy. And, although nearly all of the 50 states, Puerto Rico, and the District of Columbia submitted multiple amendment requests sometime in the past year, a few have not submitted any requests at all.

A number of factors may have contributed to the large number of amendment requests in 2003-04, such as the high stakes consequences of the NCLB accountability requirements, recently announced federal flexibility regarding several provisions of the requirements, delays in the provision of regulations and guidance from ED, exchange of information among states on elements of their plans, and the initial release of 2002-03 AYP results along with projections for subsequent years. Decisions resulting from the 2003-04 round of accountability plan amendments are also likely to trigger additional amendment requests next year or in following years.

As noted above, a few states did not submit amendments to their plans as they now stand. These states may be waiting for political events to unfold or for decisions within their own jurisdictions, conducting additional analyses to support their request, or have local factors related to their plan or planned changes that prevented a request this year.

Finally, this paper addresses publicly available state requests for accountability amendments. Some states may have requested amendments without making these requests public; thus, it is possible that a number of requests have been made that could not be captured here.

Timeline of Significant Events Related to State Accountability Plans

- The NCLB Act Enacted (January 2002)
- Standards & Assessment Regulations Issued (July 2002)
- Accountability Regulations Issued (early December 2002)
- State Meetings with ED Officials Begin (December 2002)
- ED Releases Accountability Workbooks (late December 2002)
- Draft State Accountability Workbooks Due to ED (January 31, 2003)
- Peer Reviews of State Accountability Plans (January through April 2003)
- All State Accountability Plans Tentatively Approved (by early June 2003)
- CCSSO Publishes Paper on State Plans and Approval Decisions (July 2003)
- States Begin Submitting Plan Amendments to ED (Fall 2003-spring 04)
- Regulations Regarding Achievement of Severely Cognitively Disabled Students Issued (December 9, 2003)
- Simon Letter to Chiefs Setting April 1 for Plan Amendments (February 2, 2004)
- Secretary's Letter to Chiefs Announcing New Flexibility Related to LEP Students (February 20, 2004)
- Secretary's Letter to Chiefs Announcing New Flexibility Related to SWDs (March 2, 2004)
- State Amendments Due to ED for Decisions Applied in 2003-04 AYP Determinations (April 1, 2004)
- ED Response Letters Begin (May 2004)

Section II: Key Issues in States' Amendments to Their NCLB Accountability Plans in 2003-04

The central issues addressed in states' educational accountability workbook amendments and described in Part II of this document are 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

About one quarter of the states seeking amendments to their NCLB accountability plans submitted requests related to their standards and assessments. These requests fall into the following categories (some states requested amendments in more than one category):

- Changes to the test or grade level that yields scores used in AYP analyses, as indicated in the original plan;
- "Banking" of test results for use in AYP determinations when students have multiple opportunities to take the test used for AYP purposes;
- Use of out-of-level tests (also termed instructional level tests) or other alternate assessments for students with disabilities or students with limited English proficiency; and
- Adjustment of the scores students must achieve to be considered proficient.

When interpreting the information in this section of this paper, it is important to note that all states, the District of Columbia, and Puerto Rico will be required to submit evidence pertaining to their standards and assessments by the spring of 2006 for a separate federal review. This review will encompass new standards and assessments as well as standards and assessments that were reviewed under the Improving America's Schools Act of 1994 if their results will be used in the NCLB-related AYP determinations. Some states may be waiting to request accountability amendments until the completion of this next round of standards and assessments reviews.

Changes to the Test or Grade Level Used for AYP

Most states will be making significant changes to their assessment systems over the next few years, but, so far, only a few have addressed these in their requests for amendments. In some cases, states simply indicated a change in the test that will yield the scores used for AYP purposes or to meet other NCLB requirements. For example, Alabama was approved for a switch to the English Language Development Assessment⁴ (ELDA) for measuring English proficiency. Another state requested a similar switch for one of its academic assessments.

Tennessee, which already tests in grades 3-8 and originally planned to use assessment results for all of these grades in AYP analyses, asked to use results from only a subset of these grades for the years prior to 2005-06 (the year in which assessment in each of these grades is required by NCLB). Some arguments for this request pointed to a misalignment between uses of the data for AYP and the uses for which the state had developed its current assessments. Fairness was also at issue: states that already test in these grades may need the same amount of time as other states (i.e., until the 2005-06 school year) to prepare for the full testing requirement. As of August 2004, ED had not yet provided a response to these requests.

⁴See http://www.ccsso.org/projects/SCASS/Projects/Assessing_Limited_English_Proficient_Students.

Similarly, other states indicated that a set of tests or the policies associated with their implementation would not take effect prior to 2005-06 as had been anticipated. Therefore, the source of the assessment data used for AYP determinations would be somewhat different than originally described for 2003-04 and 2004-05. In fall 2003, **Maine** proposed the use of results from local assessments for AYP purposes. However, the state later withdrew this request after it determined that the federal assessment and accountability requirements might prove overwhelmingly burdensome and intrusive for its instructionally-focused local assessment system. **Maine** will continue its support for the local assessments but will also develop statewide assessments for use in AYP measures.

Banking of Test Results

A number of states received approval for the practice of “banking” test results when students have two or more opportunities to take a test that yields scores used in AYP analyses. The version of this practice that has evolved into what is now 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. A state may allow students multiple opportunities to take the test before this point. That is, if the test is based on standards relevant to the end of 11th grade, test administrations before this time may be considered “early” and the administration in the spring of 11th grade “official.” If a student takes the test before the official administration and achieves an acceptable score (i.e., proficient), that score is “put into the bank.” If that student scores below the acceptable level, the score is not banked and the student must take the test again. The re-testing student may continue to test as often the state allows, but the scores will not count for AYP judgments until he or she passes or until the official administration, whichever occurs first. States using this strategy more often do so at the high school level when the test is a requirement for graduation. States may bank for participation rate determination just as for performance: only passing attempts count prior to the official administration.

Oregon’s original accountability plan included a form of score banking that allowed credit for performance on computer-adaptive tests administered on-demand during specified testing windows each year, including multiple opportunities to re-test when the passing standard is not met. ED rejected this strategy during the initial reviews in 2003, citing the NCLB regulation (§200.20(c)(3)) requiring states to use the first score a student obtains in their AYP calculations and the lack of a clear way of calculating a participation rate. The “first score” requirement described in §200.20(c)(3) had no basis under NCLB and ED later reversed its position (Simon, January 26, 2004) in response to Oregon’s December, 2003, request for amendment. Thus, **Oregon** is now allowed to implement its original, previously-denied banking strategy and will use the spring administration as the “official” one for AYP purposes.

In spring 2004, several other states requested amendments to apply banking rules. At least one would like districts to choose when students have the opportunity to take a test, as long as it is administered by the end of 11th grade. Another state has asked for approval of its plan to use results from re-testing opportunities that occur after the official administration of a test; the re-testing students would be administered new test items covering only the standards where he or she had not previously met the passing criteria.

One other state has requested permission to use results from all grades eligible to take the test rather than results from only the official administration in AYP calculations for its districts. This state administers a test based on high school standards that students may begin attempting while in middle school. The state had initially planned to start using the results from only the highest grade level at which the test could be taken, then use results from the highest two grade levels in the following year, the highest three grade levels the following year, and so on until results from all eligible grade levels would be included. The state would like to amend its plan so that results from all grade levels are included in the district and state AYP analyses each year beginning immediately; for school-level AYP, only scores for high school students would be counted. It is not clear from the documentation how the state intends to calculate participation rate.

Out-of-Level Tests or Other Alternate Assessments

Out-of-level testing involves the administration of a test form based on one grade level's content expectations to a student who is enrolled at another, usually higher, grade level. Under regulations promulgated on December 9, 2003, and consistent with Secretary Paige's March 2, 2004 letter, out-of-level testing is effectively now defined by ED as a test based on alternate academic content and student achievement standards for severely cognitively disabled students and subject to the law's 1 percent cap restrictions for considering student performance as proficient in AYP determinations. Under these conditions, such tests may be used as part of a state's assessment and educational accountability systems.

During the initial review and approval cycle in early 2003, no state was allowed to use results from out-of-level testing in AYP calculations⁵; however, this position changed over time and by June 2003, ED was approving the use of out-of-level assessments for students with disabilities (SWDs) to make AYP determinations within a narrow range of parameters (see Erpenbach, Forte Fast, & Potts, 2003). In these instances, ED referred to out-of-level assessments as "instructional level assessments" (Paige, June 27, 2003). The use of out-of-level testing continued to be prohibited except in cases where it was argued that out-of-level results are considered necessary for instructional purposes and the student also takes an on-grade test for AYP purposes. Prior to this time, states had not been allowed to use the results of out-of-level testing in AYP calculations because these tests were, by definition, not aligned to on-grade expectations.

One state requested an amendment allowing out-of-level tests to be offered as an option for severely cognitively disabled students. Four others requested approval for the use of out-of-level tests as one of the options available for including students with non-severe disabilities in the assessment system. ED responded in one of these cases shortly after issuing the final regulations on the inclusion of students with the most severe cognitive disabilities in Title I assessment (*Federal Register*, December 9, 2003), specifically rejecting the use of proficient scores from out-of-level tests in AYP calculations unless the scores are included in a 1 percent cap:

The Department has issued a final regulation that permits a State to use alternate achievement standards to evaluate the performance of students with the most significant cognitive disabilities, and to give equal weight to proficient achievement based on the alternate standards in calculating school, district, and State AYP, provided that the number of proficient scores based on the alternate achievement standards does not exceed 1.0 percent of all students in the grades tested at the State or LEA level. Any use of out-of-level assessments for AYP purposes would be subject to this regulation. The final rule does not alter the IEP team's role in making individual decisions about how to assess a child. Instead, it restricts, solely for purposes of calculating AYP, the number of scores that can be counted as proficient based on alternate achievement standards. Achievement standards associated with out-of-level assessments may meet the requirements of this regulation only if they are aligned with the State's academic content standards, promote access to the general curriculum, and reflect professional judgment of the highest achievement standards possible. The results from out-of-level assessments must be included within the 1.0 percent cap for the purposes of calculating AYP, because the achievement standards associated with the content measured by out-of-level assessments are clearly different from the achievement standards in the target grade (Tomalis, December 31, 2003).

⁵ In promulgating regulations for NCLB standards and assessment provisions, ED stated, "Although out-of-level tests, for example, may provide instructional information about a student's progress, they are not an acceptable means to meet the State's assessment requirements under §§200.2 and 200.6 or the accountability requirements of the NCLB Act." (*Federal Register*, July 5, 2002, pp. 45044-45045).

None of the other three states that have publicly requested permission to continue using out-of-level tests has received a public response from ED; however, sources in at least one of these states suggest that ED may passively allow the use of out-of-level tests for AYP purposes simply by “looking the other way” until this state undergoes the Peer Review of its standards and assessments sometime in the next two years. This state, and at least one of the others, cites the need for more time to develop and implement other alternate assessments for students with non-severe disabilities.

In early 2004, **West Virginia** proposed the use of “an equivalent” form of its statewide academic assessment, WESTEST, that would be aligned to grade-level standards but would be more accessible for students with cognitive impairments. The state does not intend for the results from this assessment to be included in the 1 percent cap because it is to be based on grade-level expectations; therefore, there would be no limit on the number of students who could score at or above the proficient level and have their scores included in AYP determinations (provided that the assessment would meet all of the relevant NCLB Act requirements). Although approval of the assessment itself could occur only as part of the upcoming standards and assessment reviews, ED did respond to this basic request signaling a willingness to consider the state’s proposal:

This approach would be in alignment with the statute and regulations if the ‘equivalent’ assessment is aligned to grade-level achievement standards and content standards. In such an instance, there would be no limit on the number of proficient scores that could be included in AYP assessments (Simon, March 4, 2004).

Several other states have made requests to use the results of new alternate assessments for severely cognitively impaired students in their AYP calculations. Those that have received approval to date include **Alabama**, **Indiana**, and **Nebraska**. **Maine** has specifically been asked to ensure that the results from its new alternate assessment are included in its AYP calculations, although it is not clear that the state ever asked to do otherwise. (See the “*Inclusion*” section of this paper for additional information on these issues.)

Adjustment of Scores Designating Proficiency (“Cut Scores”)

NCLB requires states to develop student academic achievement standards (previously called performance standards) that define at least three levels of performance. These levels must include one that corresponds to grade-level proficiency (proficient), at least one level above proficiency (advanced), and at least one level below proficiency (basic). As part of the development process, states must identify the test scores that differentiate between these levels, referred to as “cut scores.”

Although the definition of “proficient,” including the score a student must achieve on a test in order to be classified as proficient, has much to do with the meaning of AYP, this has been the subject of only three amendment requests to date. Two states requested amendments reflecting changes in their student academic achievement standards due to the introduction of a new assessment. One other state made an adjustment in the cut score separating its proficient level from the level below it on an existing test. This adjustment was described as the correction of an error that had been made during the original standards-setting process. So far, no state has requested an amendment indicating plans to revisit its student academic achievement standards in general.

Adequate Yearly Progress (AYP) Model

A state’s AYP model defines how it actually calculates adequate yearly progress for schools and districts. This model must address questions such as, “Which students are included in the denominator when you calculate the participation rate?” and “How does the state address the reliability of its AYP decisions?”

Over half the states that submitted amendments requested some change to their AYP model. These requests may be categorized as follows:

- AYP Application;
- AYP Indicators;
- Annual Measurable Objectives, Intermediate Goals, and Trajectories;
- Strategies Intended to Enhance Reliability.

AYP Application

Several states requested permission to make changes in how and to whom AYP decision rules would be applied. Such requests related to how schools operating Targeted Assistance Title I programs calculate AYP, general triggers for identification, differential rules for identifying districts for improvement, and the retroactive application of approved amendments.

Rules for Targeted Assistance Title I Schools

For the first 20 years of its implementation, the Elementary and Secondary Education Act (ESEA) only allowed schools to operate Title I programs that targeted academic assistance to specific students (commonly referred to as “Targeted Assistance Programs”). With the Hawkins-Stafford amendments to ESEA in 1988, schools in which a significant proportion of the students come from low-socioeconomic backgrounds were given another option for providing Title I services. Instead of identifying specific students for service as part of a Targeted Assistance program (see sec. 1115 of Title I), these high-need schools could be granted permission to combine their Title I funds with dollars from other funding sources in order to implement Schoolwide services designed to improve the quality of education for all students in the school. Since these amendments were adopted, more and more schools have taken advantage of this flexibility in the use of their Title I funds, so that most high-need Title I schools now operate Schoolwide programs (see sec. 1114 of Title I) rather than Targeted Assistance programs.

Under section 1116(b)(1)(D), AYP calculations for schools operating Targeted Assistance Title I programs may be based on the progress of all students in the school or on only students who are served—or are eligible to be served—under Title I⁶. No states approved in the initial round of accountability workbook submissions appear to have chosen the latter approach.

North Carolina became one of the first states to amend its accountability plan so that some Local Educational Agencies (LEAs)—more commonly referred to as school districts—may chose to calculate AYP for schools that operate Targeted Assistance programs (as opposed to Schoolwide programs) based only on those students served or eligible to be served through Title I. ED approved this request but asked **North Carolina** to “conduct this analysis in alignment with ED guidance on this issue.” **North Carolina** will not allow districts to change a school’s program designation (e.g., from Schoolwide to Targeted Assistance) after the start of a school year. In addition, districts may choose to calculate AYP based on results for students who are served by Title I only if that group meets the state’s minimum “n” of 40. **North Carolina** subsequently confirmed with ED that states could exercise this approach every year and not just the year after which an AYP target has been missed.

At least one other state has submitted a similar request but as of August 2004 had not received its response from ED.

⁶ Title I funds are allocated to states, school districts, and schools on low income statistics; however, children eligible to be served under Title I programs are those a school identifies as “failing or most at risk of failing, to meet the State’s challenging student academic achievement standards” (sec. 1115(b)(1)). In other words, the students eligible to be served are not always those students who generate the funding.

Triggers for Identification

At the most basic level, NCLB requires that schools or districts that miss their AYP targets in two consecutive years be identified for improvement. There are a number of possible permutations of this basic idea, and some states requested amendments related to the specific set of conditions that result in identification for improvement. **Arkansas** was approved (September 2003) to conduct comparisons of current-year assessment results with the average of the most recent three years' results (including the current year results) and to use the most favorable data in making AYP determinations. ED required the state to (1) submit impact data on the number of schools and districts making AYP because this procedure was applied and (2) assure that the same procedures will be applied to each subgroup within a school. **Washington** was approved (June 18, 2004) to "move towards averaging data across grade levels, and will average data over the past two years for AYP determinations."

In early 2003, ED's position was that a school or district missing any AYP measure for two consecutive years would result in identification for improvement. For example, a school in which a student group failed to meet the annual measurable objective (AMO) for mathematics in one year and a different group at the same school failed to meet its other academic indicator (OAI) the following year would be identified for improvement. That position evolved in subsequent approval decisions such that states may now restrict AYP determinations to the same subject (an amendment approved for **North Carolina** in May 2004 and **Illinois** in June 2004) or the same OAI (an amendment approved for **Oregon** in May 2004 together with its request for same subject determinations). ED, however, continues to steadfastly maintain that AYP determinations may not be restricted to the same subgroups within the subject or OAI; a subject that continues to be debated nationally given differing readings of the statutes and regulations among the states and other interested parties.

ED's rationale was included in a December 31, 2003, letter from Acting Assistant Secretary of Elementary and Secondary Education Ronald J. Tomalis to **West Virginia** State Superintendent of Public Instruction David Stewart in response to the state's proposed change in how AYP would be determined (**Utah** was another state that sought to base the identification for improvement on the failure of the same subgroup in the same subject or OAI.):

West Virginia proposes to identify schools and districts for improvement only if the same subgroup did not make AYP in the same subject or indicator for two consecutive years. This identification procedure is not consistent with the statute or regulations. As you know, there is flexibility to permit States to determine schools in need of improvement on the basis of not making AYP in the same subject for two consecutive years. This flexibility stems from other provisions in the statute that treat reading and mathematics independently—e.g., separate starting points and annual measurable objectives. These provisions recognize that student achievement in reading and mathematics in a State may be starting at very different points and thus the State would need to establish different trajectories to attain 100 percent proficiency. Concomitantly, it made sense to permit a State to identify schools in need of improvement based on not making AYP for two years in each subject. Subgroups, on the other hand, are not treated differently in the statute and thus the statute does not support similar flexibility to identify schools in need of improvement on the basis of "same subgroup" performance for two consecutive years. Moreover, your proposal is inconsistent with the statute's accountability provisions, which require that each subgroup meet the State's annual objectives in each subject each year. 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 elements that need to be improved. A school that is identified for improvement [the statute refers to identifying schools for improvement] should look to specific instructional remedies in the subject, other indicator, or participation rate.

A New Approach to Identifying Districts for Improvement

In one of the more surprising developments of the 2004 amendment process, at least eight states have received approvals to modify the method by which school districts missing AYP targets are identified for improvement. Prior to these amendments, all states calculated district AYP based on results for all students in the tested grades (or all students in the tested grades in Title I schools) in the district. **Arkansas, Delaware, Iowa, Maine, North Carolina, Ohio, Tennessee, and West Virginia** have now been granted permission to calculate AYP separately for up to three grade spans. That is, they can calculate AYP for elementary schools, AYP for middle schools, and AYP for high schools. For a district to be identified for improvement, it would have to miss AYP at each of these levels two years in a row. Further, these states can consider AYP separately for reading and mathematics such that a district would have to miss AYP in reading (for example) at all three levels two years in a row to be identified for improvement. ED's response letters to **Delaware, North Carolina, and Ohio** noted that the state should "1) monitor districts that have not made AYP in one grade span but have not been identified for improvement to ensure they are making the necessary curricular and instructional changes to improve achievement, and 2) take steps to ensure supplemental [educational] services are available to eligible students from a variety of providers throughout the state (including in LEAs that have not been identified for improvement but that have schools that have been in improvement for more than one year)."

The authors are aware of a number of other states that are preparing to request similar amendments and note that this strategy may result in many fewer districts being identified for improvement shortly before the federal elections in the fall of 2004. In a small state preparing to submit an amendment, for example, the number of school districts that would be identified for improvement based on 2003-04 determinations would drop from 100 to 15. In **Tennessee**, 92 percent of the school districts did not make AYP last year. The approved amendment dropped the proportion to 65 percent (Olson, July 14, 2004).

Retroactive Application

While most requests for amendments were specific to a particular component of a state's accountability plan, some states submitted requests for more general flexibility. **North Carolina** requested retroactivity in the application of approved amendments to AYP calculations for the 2002-03 school year. Further, the state requested that in the event that ED did not approve such retroactivity, it be allowed to treat the 2002-03 school year as a pilot year and the 2003-04 as the first year for official AYP reporting. Both of these requests were primarily based on **North Carolina's** desire to ensure consistency and comparability of the 2002-03 AYP decisions in the state with the 2003-04 decisions. ED denied both parts of this request:

It would be inconsistent with Congressional intent, legal precedent, and the facts of implementation to allow the retroactive application of policies and decisions, even those made by a State. NCLB accountability requires that schools are only identified for improvement after not making AYP for two consecutive years. Schools that did not meet AYP last year because of state decisions about AYP, and then make AYP this year under new state rules, will not be identified for improvement. Thus, allowing for the retroactive application of policies and decisions made by States would result in no practical difference, and would also result in great confusion in the field about what rules applied when and how. To do otherwise would render state accountability systems ineffective and indefinable; consistency would be lost; and schools, teachers, parents and students would face a moving target that provided very little, if any, standard by which to measure progress and success (Simon, May 20, 2004).

In another part of its response to **North Carolina**, ED made reference to a letter Secretary Paige sent to Senator Edward M. Kennedy on April 27, 2004 (Paige, April 27, 2004). The letter was in response to one Senator Kennedy⁷ had written to Secretary Paige on March 24, 2004, requesting that states be allowed to

apply the “regulations” related to the inclusion of students with severe cognitive disabilities (*Federal Register*, December 9, 2003) and students with limited English proficiency (February, 2004) retroactively to AYP calculations for the 2002-03 school year. Senator Kennedy noted that these regulations had not been released in a timely manner, suggesting that it would be appropriate to allow states to apply the new regulations consistently to NCLB AYP decisions for all years. In his response, Secretary Paige rejected the timeliness argument and highlighted the lack of precedent for retroactive implementation of federal policies. He went on to indicate that no school could be identified without missing AYP twice consecutively and, therefore, could not be identified in the absence of the new flexibility that was available for the second AYP year.

Although not mentioned in **North Carolina’s** request for amendments or Senator Kennedy’s letter, provisions of Part D of Title IX of NCLB (section 9401) do appear to permit the Secretary to approve retroactive application of amendments to a state accountability plan through the law’s provisions for waivers of statutory and regulatory requirements. Retroactive application in this case is not included among the law’s restrictions on waivers (section 9401(c)) nor is it prohibited under section 1111(g)(1)(B) pertaining to the extension of deadlines for state assessment and accountability systems.

In a very different approach to retroactive applications in AYP decision-making, **Kentucky** incorporated in its 2002-03 accountability determinations changes that were only later submitted to ED for its review. The changes were eventually approved and the state’s 2002-03 AYP decisions were allowed to stand with the following comment the approval letter from ED:

I also understand that Kentucky made adequate yearly progress (AYP) determinations for the 2002-2003 school year using several of the amendments listed . . . that had been submitted to the Department, but not yet approved. This unacceptable action impedes the Secretary’s duty to ensure that States are implementing accountability plans that meet the requirements of Title I, Part A of NCLB. The Secretary takes this duty seriously and is inclined to withhold Title I funds from any State that does not fulfill its assurance of implementing its approved plan. By not using its approved plan to make AYP determinations for school year 2002-2003, Kentucky is subject to the loss of federal funding. In this one instance, however, the Secretary will not withhold Kentucky’s federal funds only because the criteria it used to make AYP determinations are compliant with NCLB. In the future, do not deviate from your amended plan as approved by the Department, lest you be subject to financial consequences (Simon, February 2, 2004).

ED’s position with respect to retroactive application of approved changes to how a state makes AYP determination is somewhat perplexing. The law seems to permit just such an action, and ED decided not to penalize **Kentucky** for effectively making retroactive applications because the applications were “compliant with NCLB.” In almost all cases, states sought changes to their AYP models—and wanted to apply them retroactively—because ED’s related approval decisions changed in so many areas as the 2003 Peer Reviews progressed. Several AYP components originally denied were later approved. ED also made a number of “surprise” approvals and later described those as evidencing new flexibility around implementation of the law. Further, some states have challenged the part of ED’s rationale that making the amendments retroactive would result in no practical difference, and would also result in great confusion in the field, believing that these concerns are erroneous and best left to the states themselves to decide.

⁷ Thirty other members of Congress co-signed this letter. Technically, additional regulations pertaining to LEP students have not been issued. The Secretary did issue a letter to states (February 20, 2004) setting forth additional flexibility under the law related to this subgroup. The Secretary issued two additional flexibility letters on March 2, 2004 (regarding SWDs) and March 29, 2004 (regarding participation rates).

AYP Indicators

Most states requested amendments related to one or more of the five indicators required for calculating AYP. These include percent proficient for reading or language arts, percent proficient for mathematics, participation rate for reading or language arts, participation rate for mathematics, and at least one other academic indicator (OAI). Graduation rate must be included as an OAI at the high school level.

Only one state appears to have made an amendment request that would apply to the calculation of all AYP indicators. **Texas** received approval to “apply standard rounding rules to all aspects of AYP elements including performance targets, performance rates, participation targets, and participation rates.” The state will round decimals of 0.5 or higher to the “next highest integer” and drop decimals of .4 or lower. Interestingly, a state reported that ED had informed it that the department would approve rounding rules in which all decimals are rounded up to the next whole number, e.g., a participation rate of 94.1 could be rounded to 95.

Percent Proficient

Very few states requested amendments related to the calculation of the percent proficient. It is worth noting that states were not required to provide the algorithm they use or even to describe this algorithm they use in calculating percent proficient in their accountability workbooks. Some states use all students enrolled for a full academic year (FAY) in the denominator and others use only those students who have been enrolled for a full academic year and actually took the test. In the June 2003 CCSSO paper on the initial accountability plans, we reported that **Maryland** and **Georgia** were the first states allowed to calculate percent proficient using the number tested in the denominator (Erpenbach, Forte Fast, & Potts, 2003, p. 15). In an un-circulated 2003 Policy Brief, ED, commenting on the calculation of participation rate, noted that, “The State may, but is not required to, assign a score of ‘non-proficient’ for any students who do not participate in the assessment. This policy requires a State to count as a ‘participant’ only those students who actually complete some portion of the assessment.” In at least the **Maryland** and **Georgia** cases, this was interpreted to mean that students who do not participate in the assessments do not have to be considered non-proficient when calculating the percent proficient. This essentially removes these students from the denominator for that statistic. It is possible that other states may be calculating percent proficient with only tested students in the denominator but have not made this public.

Most of the amendment requests in this area related to changes to their definition of full academic year. States are required to include in the calculation of percent proficient only scores for students who have attended the school for at least a full academic year. **Oregon** received approval for defining full academic year as enrollment for at least one-half of the instructional days prior to May 1 of each year. One other state would like to define full academic year as enrollment for at least 173 days in the current school year; two others would like to consider students to have been enrolled for a full academic year if they are enrolled at a specific point in the fall and also enrolled at the time of the test. ED approved (July 1, 2004) **Kentucky’s** new definition of a full academic year as enrollment for at least 100 instructional days counted from the first instructional day to the first day of the state testing window.

Wyoming received approval to “use up to two years of achievement data when making AYP determinations.” The state will make a secondary examination based on averaged data whenever a school or district does not make AYP on the basis of current year results. **Iowa** received approval (June 17, 2004) to “clarify its definition of full academic year for students with IEPs who participate in the alternate assessment. Since the Iowa Alternate Assessment is a collection of evidence over time, the specifications need to indicate the length of time over which evidence is collected, and what happens to the evidence.” Effectively, then, full academic year will be defined for some SWDs as being more than one school or calendar year.

One other state has requested permission to use a growth model in lieu of percent proficient; this state did not describe the particular model it would like to use.

Participation Rate

Of all the changes states requested to their accountability designs, calculation of participation rate received far and away the greatest attention. Clearly, the decision ED Secretary Rod Paige announced on March 29, 2004, to provide increased flexibility with regard to how states may calculate participation rate in making AYP determinations was a welcome change in an area proving to be troublesome for many states.

While the number of schools missing solely the participation rate requirement since the enactment of NCLB is unknown, it appears to be considerable. Robelen (*Education Week*, 2004) profiled the experience of several states and school districts, concluding, “In many cases across the country, schools just narrowly missed the 95 percent participation rate last year.” In **California**, for example, nearly two-thirds of the high schools missed only this AYP requirement. **Texas** initially identified about 1,000 schools as not making AYP in 2002-03 with participation rate being the leading cause. Earlier, Pennsylvania, in handling hundreds of appeals from schools identified for improvement, decided to round up test-participation rates of 94.5 or above (Gewertz, 2003).

Policies for Calculating Participation Rate (March 29, 2004)

- A state may use data from the previous one or two years to average participation rate determinations for a school or subgroup.
- Students unable to take tests because of a unique, significant medical emergency do not have to be counted in calculating participation rate.

Originally, participation rate was to be calculated using the number of students enrolled at the time of testing as the denominator and the number of students who participated in the assessments in the numerator. This statistic was to be calculated independently for each school year. No enrolled students were to be excluded from the denominator. That is, prior to releasing this guidance, ED allowed no exemptions to testing and recommended that states handle exemptions due to medical emergencies as part of the appeals process under sec. 1116(b)(2) for schools and sec. 1116(c)(5) for LEAs.

Each of the states that asked to apply the new flexibility to the AYP analyses for 2003-04 school year has received approval from ED—as long as the request fits within the parameters of the guidance.

It should be noted that in December 2003, ED rejected **West Virginia’s** request to develop state criteria that would allow for limited exemptions of students “when an unexpected medical or psychological condition prohibits participation.” ED indicated that that the state could not establish “preemptive criteria for such exclusions” but may consider them on a case-by-case basis through the appeals process.

As an example of how states have incorporated the new guidance into their plans, **Ohio** plans to use two-year weighted averages for calculating participation rates. In calculating “weighted” averages, the state will take into account the actual denominators and numerators for each year and will not simply average the participation rates from the two years. **North Carolina** is another state that will adopt multi-year averaging for this AYP requirement.

Minimum “n” for Participation Rate Calculations

Requests that were not related to the new guidance included one for an increase in the minimum “n” (see the “*Strategies Intended to Enhance Reliability*” section of this paper for a definition of minimum “n”) for participation rate from 20 to 40. **Kentucky** received approval (July 1, 2004) to revise its minimum “n” for participation rate from 10 students per grade and 30 students per school to 10 students per grade and 60

students per school or 15 percent of the school population. In other words, Kentucky will not calculate participation rates for any group with fewer than ten students per grade and 60 students per school unless that group represents at least 15 percent of the enrollment at the school.

Two other states would like to adjust the participation rate target such that schools and subgroups with fewer than five students not tested, regardless of enrollment, would be considered to have met the participation rate requirements. A variety of strategies to reduce the impact of small numbers of non-tested students in small schools were included in states' original accountability plans; these "fewer than five" not tested requests for amendment may provide a simpler option if approved.

Students "Opted-Out" of Testing

In some states, parents are allowed by statute or regulations to exclude their children from state testing programs. These "opt-out" provisions can mean that a school could miss the participation rate target even if every other student participated. To protect schools in these situations, one state requested approval to omit opted-out students from the denominator of the participation rate. Parents in this state may opt their children out of assessments, although only a few do so each year. ED denied the request verbally, but did not mention this request in its response letter, and has not subsequently provided a written rationale or explanation for its non-response. Instead, Secretary Paige may have been referring to this issue when he commented in the approval letter, "those changes that are aligned with NCLB are now included in an amended State accountability plan."

Another state would like to use a more complex strategy. If a group/school/district did not meet its percent proficient target, then it did not make AYP regardless of participation rate; however, if a student group, school, or district met the percent proficient target and tested at least 90 percent of the enrolled students, the number of students opted-out would be compared with the number needed to raise the participation rate to 95 percent. If the number of opted-out students equals or exceeds the number needed to make 95 percent participation, then the participation rate target is considered met and the percent proficient is recalculated with the opted-out students counting as not proficient. The recalculated percent proficient is then compared to the target. This approach would seem to be inconsistent with ED's policy on not counting as "non-proficient" students who do not participate in state assessments (see above under "*Percent Proficient*").

Statistical Tests Applied to Participation Rate Determinations

The 2003 CCSSO report (Erpenbach, Forte Fast, & Potts, 2003) documented that ED's perspective on the application of statistical tests for the purpose of enhancing the reliability of AYP determinations changed over time. Initially, ED indicated that statistical tests such as confidence intervals could not be used with "count" indicators⁸ (i.e., participation rate, attendance rate, and graduation rate). They also could not be used in safe harbor⁹ reviews (application of confidence intervals to safe harbor is described in a subsequent section of this paper.). Some state accountability designs that were approved late in the review cycle in 2003, however, included the use of confidence intervals with "count" items. For example, **North Dakota** was approved for the use of a 95 percent confidence interval for participation rate calculations. Other states noted this approval, especially if they had proposed this method earlier or had been initially interested in it but somewhat dissuaded by ED's initial decisions.

⁸ "Count" indicators are rates—that is, the result of dividing the number of students in a group that meet certain criteria by the total number of students in the group. Percent proficient is also a rate, but apparently is considered by ED to be a different category of indicator.

⁹ See Appendix B in Marion, S., White, C., Carlson, D., Erpenbach, W. J., Rabinowitz, S. & Sheinker, J. (2002). *Making valid and reliable decisions in determining adequate yearly progress*. Washington, DC: Council of Chief State School Officers (available at <http://www.ccsso.org/content/pdfs/AYPpaper.pdf>), for a review of the safe harbor provisions and their application.

On initial review, it appears that ED continues to deny the use of a statistical test with “count” elements. Both **Tennessee’s** request to use confidence intervals in attendance and graduation rate calculations and another state’s request to use them in participation rate calculations were denied. Many states, however, appear to have approval to use confidence intervals on such indicators; ED approved their proposed amendments to use confidence intervals for “AYP calculations” or “AYP decisions” without reference to (without exclusion of) specific AYP elements.

Other Approaches to Calculating Participation Rate

Kansas received approval for its request to use the number of test answer documents as the denominator in its participation rate calculation because the state does not currently have a student-level database. A number of other states already use this strategy; all are required to enforce the return of answer documents for every enrolled student, whether a student tests or not. Another state allows students to retest in the areas where they did not achieve a passing score on the first attempt and would like to base participation rates only on first-time test takers.

Other Academic Indicators

Requests for amendments to other academic indicators (OAIs) relate either to those used at the elementary and middle school levels or to graduation rate (required) at the high school level.

Other Academic Indicators at the Elementary and Middle School Levels

Several states requested changes to their other academic indicators at the elementary and middle school levels. The most common change seemed to involve a shift from requiring schools and districts to meet or exceed a specific target to one of making progress toward the target (**Hawaii** being an exception). For example, **Tennessee** changed its AYP requirement that schools and districts meet or exceed an attendance rate of 93 percent at the elementary school level and a graduation rate of 90 percent at the high school level to also allow for any degree of progress over rates from the previous year. This state also requested approval to employ a 95 percent confidence interval with these rates, which ED denied. Schools and districts above the targets are not required to make further progress.

Pennsylvania changed its AYP targets for attendance and graduation rate from 95 percent each to 90 percent and 80 percent, respectively, to permit continued progress toward these targets for schools and districts below them. **Kansas** dropped an attendance rate requirement at the high school level that it had originally applied for AYP determinations in addition to the graduation rate requirement at this level. Other states have requested a simple change to the status target or the use of a single year or the average of three years of data, whichever provides the greater benefit to the school.

Several other states requested permission to replace the indicator described in their original accountability plans with another indicator. **Hawaii** will use a retention rate for its OAI (the percent of repeating 8th grade students) but will require that schools meet or exceed the state-set targets; not just make improvement over the previous year. **Virginia** will now allow elementary and middle schools to choose either the current OAI, attendance with a target of 94 percent, or performance on the state’s science assessment with a target of 70-percent proficient. **Kansas** will be allowed to remove attendance as an OAI for high schools.

Kentucky will be allowed to lag its OAIs by one year in order to ensure reliability of the data used. It will also be able to use its accountability index rather than its academic index for the OAI. The academic index includes only achievement information while the accountability index includes attendance and retention in the elementary school grades and attendance, retention, and school drop-outs in the middle grades. Graduation rate will remain as the only OAI used at the high school level.

Other Academic Indicators at the High School Level

States requesting amendments to the graduation rates¹⁰ asked for changes to when they are applied, how they are calculated, or the criterion to which they are compared.

Washington was approved to lower its graduation rate target from 85 percent to 66 percent based on a new analysis of more accurate data on current graduation figures, but also established 85 percent as the long-term goal for the 2013-14 school year. In situations where a student who has dropped-out of school re-enrolls and then drops-out again, **Maine** will be allowed to count only one drop-out event. In other words, the student is the unit of analyses, not the drop-out event. Another state will not use graduation rate as an OAI for any school that does not have as its mission the graduation of students. This means that, for example, alternate high schools whose mission is to return students to comprehensive high school will use something other than graduation rate as an OAI.

Delaware modified its definition of graduation rate to include “or in the timeframe specified in the IEP”¹¹ and added interim annual goals toward reaching its 2013-14 target of a 90-percent graduation rate. Schools and districts must meet the annual interim goals or maintain progress toward the 2013-14 target to make AYP. Other states modifying their graduation rate definitions to include as graduates SWDs earning diplomas within time frames specified in their IEPs included **Kansas, Kentucky, Tennessee, and Washington**.

Wyoming will be permitted to review the graduation rate for very small high schools before making AYP decisions. Small schools with less than 30 “exiters” will be examined individually to ensure a valid decision. This will include a review of three years of data to see whether positive progress is demonstrated in the graduation rate over time.

Another state asked to calculate graduation rate as a probability model, but as of August 2004 had not yet described this model in detail.

NCLB specifically states that students earning General Education Diplomas (GEDs) cannot be counted as graduates in the numerator of the graduation rate. One state, however, has requested an amendment to allow GED recipients to count in exactly this way. Another state has asked for permission to shift the calculation of its graduation rate to another, yet-to-be-determined method.

Annual Measurable Objectives, Intermediate Goals, and Trajectories

Every state is required to specify the annual performance targets and intermediate goals that, together, create a trajectory leading to all students attaining proficiency in reading or language arts and mathematics by the 2013-14 school year. “Annual measurable objectives” may be the same each year or increase each year but, like the “intermediate goals,” must increase at least every three years. The increases for intermediate goals must be uniform in size (see sec. 1111(b)(2)(G)&(H) of Title I). For example, if a state’s starting point for reading at the elementary school level was 20-percent proficient in 2001-02, it might simply divide the distance to 100 percent (80 percentage points) by five and then set targets that increase by 16 points every other year. Some states used a technique much like this one in setting their trajectories, while many others set increases for every three years in the beginning of the trajectory and for every year near the end.

¹⁰ NCLB requires states to use graduation rate as the Other Academic Indicator at this level. States may use additional OAIs to make AYP determinations but they may not be used in a compensatory manner.

¹¹ Technically, although not provided for in NCLB, regulations issued in December 2002 “opened the door” to modify the definition of graduation rate for SWDs. However, few states appear to have taken note of this until it was included in Secretary Paige’s flexibility communiqués.

Since there was not much leeway in how trajectories could be designed and many states had already taken advantage of what little flexibility was allowed, few states requested amendments related to their annual measurable objectives or trajectories. **Kentucky** received approval to set additional starting points, annual measurable objectives, and intermediate goals based on grade span configurations that would cover all of the state's school and district configurations including, for districts, K-8, 7-12, and K-12. At least one state would like to round its percent proficient and AMOs to whole numbers. **Washington** will increase its AMOs in equal increments every three years beginning in 2004-05.

Arkansas was approved in September 2003 to use a confidence interval to make AYP determinations involving AMOs and Intermediate Goals (IGs). According to the state, they will use 0.5 of a standard deviation to establish a lower range for each AMO and IG. The lower range will be re-established annually, based on most recent test results.

Strategies Intended to Enhance Reliability

Either in their original accountability plans or through the amendment process, all states have employed strategies intended to enhance the reliability of their AYP decisions. All states have identified a minimum number ("n") of cases necessary for AYP analyses (in some cases this number is, implicitly, one) and some also use statistical techniques to account for the sampling or measurement error score components. Safe harbor may also be considered as another prevalent strategy for enhancing reliability.

Minimum "N"

Using a minimum "n" strategy means that a state will only conduct AYP analyses for groups with at least a specified number of scores. Most states restrict their AYP analyses to groups of at least 30 or 40. For example, when the minimum "n" is 30, a state will not conduct AYP analyses for a school or group in which 29 or fewer scores are available for the percent proficient analysis. This often means that AYP analyses may be conducted for the school as a whole, but not for groups within the school, such as students with disabilities.

A number of states requested changes to their minimum "n." It should be noted that if the total number of scores for a school is less than the minimum "n," the state must still evaluate whether that school is making adequate yearly progress in some other way. For example, the state may, but is not required to, allow scores to be "rolled up" for two or three years if that will result in the minimum "n" being reached in order to make an AYP determination.

Arkansas will increase its minimum "n" for schools and districts from 25 to 40. **Kentucky** will increase its minimum "n" from the conjunctive minimum requirements of (1) ten students per grade and (2) 30 per school to (1) ten students per grade and (2) 60 per school. **North Carolina's** group size will change to 40 or 1 percent, whichever is greater.

Some states would like to change their overall minimum "n" while others would like to change the minimum "n" only for some student subgroups or only for school districts. In its original plan, **New Jersey** did not specify a minimum "n," but asked to amend its plan to restrict participation rate analyses to groups of at least 40. The state will use 35 as the minimum group size for SWDs and a minimum "n" of 20 combined with a misclassification rate of 5 percent (error rate) for other AYP applications.

Subgroup Minimum "N"

As noted in our June 2003 review of ED's approval decisions (Erpenbach, Forte Fast & Potts, 2003), one of the surprising approvals came in respect to the use of larger minimum "n's" for subgroups. That is, a number of states were allowed to set one minimum "n" for the school as a whole and another for the groups or

subgroups within the school. This did not escape the attention of a number of states which did not initially make such differentiations. In March 2004, for example, **Missouri** received approval for its request to increase the minimum “n” for limited English proficient (LEP) students and students with disabilities (SWDs) from 30 to 50 for accountability determinations.

Six states requested changes in the *minimum “n” for some student groups*. **Missouri** will be allowed to increase its minimum “n” for the LEP student group to 50. **Kansas** received approval in May 2004 to increase its cell size for the SWDs subgroup from 30 to 40 students with other subgroups remaining at 30. **New Jersey** will increase its minimum “n” for SWDs from 20 to 35. **South Carolina** will use a minimum “n” of 50 for SWDs and LEP students. Two other states have requested increases in the minimum “n” to 35 or 40 only for the students with disabilities group and another to 40 for both the LEP student and SWDs subgroups. The fifth state making a request in this area would like to maintain its minimum “n” of 5 but only conduct AYP analyses for groups that represent at least 15 percent of the student population. Further, this state would like to include students in no more than one special services subgroup (SES, SWD, or LEP).

District Minimum “N”

At least four states would like to use a larger minimum “n” for school districts than for schools. One has requested an increase in its minimum “n” for districts from 25 to 120; another state has asked to increase its district minimum “n” from 11 to 40. A third state currently uses a two-part minimum “n” decision rule in which AYP analyses are conducted (1) for any group of at least 100 and (2) for groups of at least 50 if those 50 students constitute at least 15 percent of the total school enrollment. This decision rule is often represented as “100/50/15%.” This state would like to retain the two-part decision rule, but increase the numbers for districts to 200 in the first part and 100 if those 100 students constitute at least 15 percent of the total district enrollment (i.e., 200/100/15%).

Another state has proposed using a 70/40/15% decision rule. (In other words, AYP analyses are conducted for groups with at least 70 and for groups with at least 40 if those 40 students constitute at least 15 percent of the total enrollment.) This state argued that larger schools and districts should not be held independently accountable based on a small number of students relative to overall enrollment. The state has also included in its proposed amendments a tiered, quantitative and qualitative review process for AYP determinations in small schools. None of these states has received a public response from ED on these requests to date.

For small schools and schools without tested grades, **North Carolina** received approval for a new process that entails “an on-site visit from the School Improvement Division.” **Texas** will offer “several approaches for AYP determinations in schools that are too small or have no students in the tested grades.”

Progressive Minimum “N”

Although ED does not appear to have approved any amendments directly permitting the use of a larger minimum “n” for school districts than for schools, at least two states have received approval for the use of strategies that will likely have the same effect. **Washington** received approval (June 16, 2004) for using what might be considered a “progressive” minimum ‘n’. In **Washington**, this means that the minimum “n” is 30 for schools and districts with fewer than 3,000 students. For schools and districts with more than 3,000 students, the minimum “n” is 1 percent of the enrollment; in other words, AYP analyses will be conducted only when the group size is at least 1 percent of the total enrollment. It should be noted that the state will also “make regular AYP determinations for small schools when there are between 10 to 29 students, with the application of a confidence interval (99%). A review of each school’s plan will be conducted “for schools smaller than 10.”

Washington will also apply the progressive strategy using a minimum “n” of 40 for LEP student and SWDs subgroups for schools and districts with fewer than 4,000 students. AYP analyses for these groups in larger

schools and districts will be restricted to cases where the number of students in the group is at least 1 percent of the enrollment.

Arkansas has also been approved to use this strategy. The state will use an “n” of 40 until enrollment exceeds 800 and then the minimum “n” will be 5 percent of daily enrollment up to 10,000 students when the minimum “n” will be capped at 500.

With regard to the aggregation of data, two states would like to combine data across two or three years if a group, school, or district does not make AYP using only the current year data. Another state wants to allow schools to choose from several options when fewer than 50 students are in a cell: use the scores as they are, use a two-year average, use a confidence interval, use results for the school serving lower grade levels from which most or all of its students come, or use the LEA score. Yet another state will no longer allow the aggregation of data across years to reach a minimum “n.”

Two states requested decreases to their minimum n’s. In both cases, the states will use a confidence interval for groups of any size. Another state plans to maintain its minimum “n” of five, but asked to count students only in the “all students” group and in not more than one program-related subgroup (scores for each student count in at least two AYP subgroups—all students and the student’s race/ethnicity category—and sometimes in several). For example, a student may have an IEP and be served in a special education program and also be an English language learner and served in an ESL or bilingual education program. This state would like to count this student’s score only in the total and the SWDs subgroup results and not in the LEP subgroup. The student’s scores would also be counted in his or her racial/ethnic subgroup. As of the printing of this report, ED had not yet responded in this case, Delaware proposed in its original accountability workbook to apportion the scores of students by subgroup membership. That is, if a student qualified in three subgroups that student’s score would be weighted by one-third in the calculation of results for each subgroup. ED denied this request.

Confidence Intervals and Standard Errors of Measurement

Many states use confidence intervals to enhance the reliability of the AYP decisions. This strategy warrants some explanation here, but readers are referred elsewhere¹² for more in-depth information about confidence intervals and their use in NCLB AYP analyses. Confidence intervals are statistical estimates of the range in which a school or district’s “true” AYP score (e.g., percent proficient) might fall, given the score it actually obtained and a few other pieces of information. The AYP score that a school or district actually obtains each year is really only an estimate of this “true” score because it is based on a sample of the students who might attend that school over time.

Because the AYP score is only an estimate, it includes some degree of error. How much error cannot be known for sure, but is expected to decrease with factors such as the number of scores on which the school or district’s AYP score is based. In other words, as the number of students’ scores on which AYP is based increases, the accuracy of the AYP score is likely to increase.

In making AYP decisions, the only known quantities are the score that the school or district earns and the target score the school or district must meet in order to have “made AYP.” Since a state knows that the earned score is only an estimate and includes some error, it must acknowledge that there is some possibility that the school or district’s “true” score is above the target even if the earned score is below the target. To protect against the identification of schools and districts that scored below the target, but whose “true” scores are above the target, some states use confidence intervals that help to account for the error in the earned scores. To use this strategy, a state would have to determine how confident it wishes to be that a school or

¹² For further information, see Hill & DePascale (2003); Linn, Baker & Betebenner (2002); and Marion, et al., (2002).

district has not met the target; most states have decided that 95 percent or 99 percent confidence is appropriate.

As they are typically used for AYP purposes, confidence intervals involve the calculation of the range above the earned score in which the state would be 95 percent or 99 percent confident that the “true” score would fall. These confidence intervals are “one-tailed” because they involve only the range above the earned score.

A third strategy meant to enhance reliability of the AYP decisions accounts for the measurement error associated with students’ test scores. Each time a student takes a test, the score he or she earns is merely an estimate of the student’s “true” score. These estimates include some degree of error. At the school or district level, this means that the AYP obtained on one testing occasion might be slightly different if the same students took the test again on another occasion (assuming they remembered nothing of the previous test questions). To account for this instability, a few states calculate the standard error of measurement and consider the school or district with an AYP score that is below, but within one or two standard errors of measurement of, the target to have made AYP.

The error associated with sampling is much larger than the error associated with the measurement (Marion et al, 2002). Most states have opted to account for sampling error and not directly for measurement error.

With regard to the use of confidence intervals—

- Several states increased their confidence intervals around percent proficient from 95 percent to 99 percent;
- Some states adopted the use of confidence intervals around percent proficient that did not include them in their initial accountability designs—generally at the 95 percent and 99 percent levels; and
- Several states added the use of confidence intervals to safe harbor reviews.

At least seven states asked for permission to use a confidence interval or use a broader confidence interval when determining whether a group, school, or district has met its percent proficient target. **Alabama** and **Kansas** (May 2004, increased from 95 percent) are among states approved to use a 99 percent confidence interval; one other state would like to use a 98 percent confidence interval. **Washington** will use a 99 percent confidence interval. **North Carolina** (approved May 2004), **Pennsylvania** (approved June 2004), and one other state asked to use 95 percent confidence intervals. The other state in this case has asked to apply the 95 percent confidence interval retroactively to the 2002-03 AYP analyses. **South Carolina** received approval to consider the standard error of measurement for the calculation of percent proficient when making AYP decisions.

Safe Harbor Reviews

The term “safe harbor” does not actually appear in the NCLB legislation or its related regulations. It was publicly adopted in 2002 to characterize a provision of the law (see Section 1111(b)(2)(I)(ii)) permitting schools and districts initially identified for improvement on the basis of missing an subject area AYP target to avoid identification if the percent of students who did not meet or exceed the proficient level of academic achievement for the year in question decreased by 10 percent from the percent not proficient the preceding school year.

References in the law that pertain to the calculation of what is known as safe harbor seem to prohibit the averaging of data across in a manner afforded under the “uniform averaging” options in Section 1111(b)(2)(J). However, at least one state received tentative approval in June 2003 to use “multiple years of data to calculate safe harbor in future years.” However, the state did not follow up on ED’s request to provide the additional information necessary to implement this procedure. Whether other states have sought

to average data across two or three years for safe harbor purposes is unknown as is the response, if any, of the U.S. Department of Education.

ED's position with regard to the use of confidence intervals in safe harbor determinations has also changed over time: states were originally not allowed to use confidence intervals for safe harbor but now are. The initial approvals were granted to **Louisiana** and **Utah**. **Louisiana** uses a 99 percent confidence interval for "safe harbor" in combination with a three-tiered improvement identification process. ED's approval letter stated that, "This procedure is allowed only in conjunction with the Louisiana growth measure, which also identifies schools for improvement if they do not make their growth target." The letter also stated that "analysis was needed regarding this particular aspect of the accountability system." The state was required to provide impact data based on its 2002-03 school year assessment results.

Utah was approved in June 2003 to use a 75 percent confidence interval in safe harbor reviews with the proviso that the state provide 2002-03 impact data as soon as it became available. ED later affirmed the state's continued use of a 75 percent confidence interval in safe harbor reviews.

Since taking their position with **Louisiana** and **Utah**, ED has also approved the use of confidence intervals for safe harbor determinations without an accompanying "growth measure," but at a lower level—75 percent. Approvals have been granted to **Alaska** (June 2004), **Kansas** (May 2004), **Maine** (October 2003), **Pennsylvania** (June 2004), and **Wyoming** (July 2004). **Pennsylvania** asked for a 95 percent confidence interval but ED allowed only 75 percent instead.

In granting approvals, ED did not provide a written rationale for setting a 75 percent ceiling (even to **Pennsylvania**, which had requested permission to use a 95 percent confidence interval) but stipulated that each state must provide impact data regarding the number of elementary, middle, and high schools that used safe harbor and, of these, the number that met AYP by invoking safe harbor and the number that still missed AYP.

Kansas did not originally conduct safe harbor reviews for small schools (those with less than 30 in the all category) as their AYP was to be determined by combining two to three years of data. According to the state, "In order to use 'safe harbor,' these combined scores needed to be compared against the previous years' scores, yet the combined scores were already included in the determination of the current year's scores. There was no comparison year that could be used that had not already been used." In order to afford small schools the safe harbor review provisions of Title I, **Kansas** received approval (May 2004) to use current year's results (even though the minimum "n" of 30 is not met) alone and compared with previous year's results.

Inclusion of All Students in Accountability

After changes in the calculation of participation rates, the two most frequently sought changes in state educational accountability workbooks were in regard to adopting new flexibility related to the assessment of students with limited English proficiency (LEP) and students with disabilities (SWDs). Almost all states seeking amendments in this area intend to take advantage of the flexibility regarding LEP students. Most decided to take advantage of the new policies and flexibility issued by the U.S. Department of Education with regard to the inclusion of SWDs in alternate assessments and accountability determinations.

The inclusion of all students, and particularly limited English proficient students and students with disabilities, in state assessment and accountability systems was emphasized in the 1994 reauthorization of the Elementary and Secondary Education Act (the Improving America's Schools Act). Among other provisions it required the disaggregated reporting of achievement results for these and other student groups. This emphasis intensified in the NCLB legislation reauthorization through the assessment, evaluation, and

AYP requirements for student subgroups. Inclusion of LEP students and SWDs in state assessment systems has posed substantial challenges for all states.

Amendments submitted by states for their accountability plans appear, in large measure, to be linked closely to three inter-connected events—(1) 2003 accountability plan review decisions made by ED, (2) amendments to regulations pertaining to SWDs, and (3) announcements by Secretary Paige of new policies and guidance earlier this year. It seems clear that the initial plan review decisions played an important role in triggering the subsequent two events. These major events/decisions are—

- Use of a larger minimum “n” for students in these two groups as compared to other student groups in making AYP determinations;
- Broad use of section 9101(25) of Title IX to include in subgroup AYP determinations LEP students receiving direct services and LEP students being monitored based on their achievement on academic assessments;
- Modification of the LEP definition to include students who are not currently receiving ESL or bilingual services but did receive them during the past one or two school years. At least two states—**California** and **South Carolina**—received approval in 2003 for this definition. The Secretary’s February 2004 flexibility letter made it clear that the test scores for students no longer meeting the state’s definition of LEP may continue to be included in the subgroup calculations for up to two additional years but do not have to be counted as LEP students for minimum “n” purposes;
- Modification of the SWDs definition to include students no longer directly served as long as their IEP calls for monitoring of progress;
- Use of alternate assessments based on alternate academic content and student achievement standards for students with the most severe cognitive disabilities (the ED flexibility letter, which contains a link to the regulations, is available at <http://www.ed.gov/admins/lead/account/csso030204.html>);
- Permission for states to apply to the Secretary “for exemptions in order to exceed slightly the 1.0 percent cap” related to SWDs taking alternate assessments aligned to alternate standards who can be counted as “proficient” in AYP calculations;
- Exclusion from AYP decisions the scores of LEP students in reading or language arts and mathematics assessments during the first year of their enrollment in United States schools. Participating students can be included in participation rate determinations (see U.S. Secretary of Education Rod Paige’s letter dated February 19, 2004, at <http://www.ed.gov/news/speeches/2004/02/02192004.html>).
- Expansion of the definition of graduation rate to recognize instances where an IEP allows students with disabilities to take additional years to graduate.

Flexibility for State Inclusion of SWDs and LEPs

Students with Disabilities: States are currently allowed to assess students with severe cognitive disabilities relative to alternate (i.e., not on grade-level) academic content and student academic achievement standards. Participation in this alternate assessment is not directly limited; however, districts can only count up to 1 percent of their total student population as proficient in relation to these alternate standards. If more students score in the proficient range, the excess over 1 percent must be counted as not proficient. Thus, ED has placed a 1 percent cap on the AYP use of scores from alternate assessments not based on grade-level expectations that can be reported as proficient in making AYP determinations. On December 9, 2003, ED released regulations permitting states to apply to the Secretary “for exemptions in order to exceed slightly the 1.0 percent cap” related to SWDs taking alternate assessments aligned to alternate standards.

Limited English Proficient Students: Exclusion from AYP decisions the scores of LEP students in reading or language arts and mathematics assessments during the first year of their enrollment in United States schools. In addition, participating students can be included in participation rate determinations (see U.S. Secretary of Education Rod Paige’s letter dated February 19, 2004).

States receiving approval to adopt the new flexibility requirements related to the assessment of LEP students include **Alabama, Alaska, Georgia, Illinois, Indiana, Kansas, Kentucky, North Carolina, Washington, Wisconsin, and Wyoming.**

Alabama, Alaska, Georgia, Illinois, Indiana, North Carolina, Washington, Wisconsin, and Wyoming were among those states receiving approval to modify their accountability systems by adopting the new flexibility related to the use of alternate assessments for severely cognitively disabled students. **Kansas** was also approved (May 2004) to exercise this flexibility but included a reallocation process to be used if a school district exceeds the 1.0 percent cap.

Alaska received approval to increase the minimum “n” (see discussion under *Strategies Intended to Enhance Reliability* for an explanation of this term) for LEP students and SWDs from 20 to 40 students for performance decisions only. Last year, 65 schools in the state did not make AYP solely because they missed the targets for SWDs (Olson, July 14, 2004). ED’s response letter to **Georgia** (June 7, 2004) noted approval for plans to “develop [in 2004-05] a system for offering additional monitoring and support services” for LEP students and SWDs. The additional services would effectively continue LEP students and SWDs transitioning out of that status in the two subgroup AYP determinations. **South Carolina** will include in the SWDs group students who are no longer directly served as long as their IEP calls for monitoring. **Kansas, Kentucky, and Tennessee** are among states now granted permission to count as graduates students completing high school with regular diplomas in the extended amount of time allowed in their IEPs.

On the other hand, ED did not go along with **Tennessee’s** proposal to increase its minimum “n” for LEP students and SWDs from 45 to 55 noting that no school or district with at least 45 SWDs made AYP last year as it was. ED also denied **Tennessee’s** request to include gifted and talented students in the SWDs subgroup. The request was based on the fact that state law places these students “under the special education umbrella” (Olson, July 14, 2004).

AYP Consequences and Reporting

For the handful of states requesting changes related to consequences and reporting, the topics of interest were the timing and degree of consequences, the provision of rewards and special recognitions, changes to demographic or program variables, and reporting timelines.

Timing and Degree of Consequences

Under NCLB, districts must provide students who attend schools that are in their first year of improvement status the option to transfer to other schools. In their second year of improvement, schools must provide supplemental educational services. **West Virginia** was among a few states that sought to change the order of these first sanctions; namely, to offer supplemental educational services rather than choice for schools in the first year of identification for improvement. ED denied the request to change the order of these sanctions, but did point out that both school choice and supplemental educational services may be offered in the first year of improvement:

The statute clearly indicates that choice must be offered to all students in schools when first identified for improvement.

The recently updated Public School Choice non-regulatory guidance provides information about how a local educational agency (LEA) could offer both choice and supplemental educational services to students enrolled schools in its first year of school improvement (<http://www.ed.gov/policy/elsec/guid/schoolchoiceguid.doc>). An LEA may give students enrolled in schools in their first year of improvement the opportunity to obtain supplemental educational services, so long as they also offer

those students the opportunity to change schools; however, because the law requires the provision of choice (but not supplemental educational services) to these students, all students who want to change schools must be able to do so, and their transportation needs must be met (subject to the 20-percent limit) before any of these students are given supplemental educational services. LEAs that offer parents of those students the option of having their child change schools or receive supplemental educational services must make it clear to the parents that, depending on the demand for choice (and the cost of transporting students to their new schools), supplemental [educational] services might or might not be provided.

In addition, if an LEA has both schools in their first year of improvement and schools in their second year of improvement or undergoing corrective action or restructuring, it must give priority for supplemental educational services to students enrolled in the schools in their second year of improvement or in corrective action or restructuring (the students who, under the statute, are entitled to be given the opportunity to receive those services) (Simon, 2004, March 4).

ED denied **Utah's** request to target school choice and supplemental educational services to "non-proficient" students. Another state asked to lag consequences by one year, so that the accountability cycle would allow for testing in the spring, reporting in the fall, and notification of parents in the following spring. This would presumably allow more adequate time to make decisions about school choice. One other state would like to tailor the degree of sanction to the level of need of the school, but did not indicate in its request how it would determine this level of need.

Rewards and Recognition

A few states asked to expand the range of rewards and recognition for schools that meet criteria beyond those required under AYP. **Alaska** (June 2004) added distinguished schools, schools that have significantly closed the achievement gap without the application of confidence intervals, and schools that have made the greatest gains in student achievement to those being rewarded under their accountability system.

North Carolina, which has integrated NCLB requirements into its existing state accountability system, added a new level of recognition, "Honor School of Excellence," for schools that make both AYP and achieve the highest ranking in the NC/ABC program. To achieve the latter, a school must have 90 percent or more of its test scores at or above the proficient level and meet its expected growth criteria under the North Carolina accountability system. **Texas** will be implementing a similar mixed classification system, whereby schools that meet state-specified criteria in addition to the AYP criteria are eligible for special recognition. **Delaware** will be using a 95 percent confidence interval for percent proficient and will now implement a new level of recognition for schools that earn scores above the AYP target (as opposed to scoring below the target but within the range of the confidence interval) and also meet other criteria.

New Jersey will set priorities for instructional intervention efforts with school districts based on the extent to which AYP was not made (June 2004).

Changes to Program or Demographic Variables

Illinois and **Maine** will be allowed to add a multi-ethnic category to their reporting and AYP groups. **Maine** will further be allowed to delete the Caucasian category, since the proportion of students in that category does not differ significantly from the "all students" group for any reporting level, and to combine the Asian and Pacific-Islander categories.

One state that does not have a statewide student-level database and does not have a student-level variable for students' socio-economic status (SES) has requested permission to substitute school-level Title I status for the student-level SES variable. In other words, the state wants to designate every student who attends a Title I school as low-SES and every student who attends a non-Title I school as high-SES. This would not allow any disaggregation by SES; districts and the state would only be able to distinguish between Title I schools and non-Title I schools.

Reporting Timeline

California and **South Carolina** were granted permission to combine their previously separate state-mandated and NCLB Act-mandated report cards. **Texas** will now align its state reporting timeline and its NCLB Act reporting timeline.

Section III: Conclusions and Observations

In our previous review of state accountability workbook plan submittals and ED's decisions (Erpenbach, Forte Fast and Potts, 2003), we summarized plan components which were consistently denied approval and those for which there were unanticipated approvals that occurred through considerable negotiation between a state and ED. Many of the decisions in this latter category were somewhat surprising to those who had been following the entire decision process. Below, we again summarize states' requests for amendments using these categories. As before, a few negotiated approvals fell into the "surprise" category. For this review, we have added a third category as well: Requests that Await Public Response

With regard to this third category, it is important for readers to remember that, very often, ED only responds publicly to some of a state's requests. Off-the-record, several states have described the following type of scenario: A state makes a formal, written request to ED to amend its accountability plan; the request includes five specific changes to the plan. ED responds publicly in writing, but only addresses four of the changes in this response. For the fifth, ED ignores the request or perhaps denies it orally over the telephone, but refuses states' requests to document this response in writing or to provide rationale for its decision. In most cases, the response letter from ED includes this statement, "Following our discussions with your staff, those changes that are aligned with NCLB are now included in an amended state accountability plan." States are thus left to conclude that ED has somehow determined that the requested change in question must not be "aligned with NCLB" but how, why, or on what basis is unknown (at least for the record). What remains unclear is whether a state must accept the oral response, this blanket statement, or whether it could have grounds to request a written response before it accepts the implied denial.

Given the interpretive caveats noted in the introduction, no overall summary indicating the number of states requesting specific amendments is provided here. Readers are encouraged to consider states' contexts when interpreting specific amendment requests. As we stated at the paper's beginning, the purpose of this paper is to present information regarding amendments sought by states to their NCLB Act accountability systems and ED's responses to those requests. Readers need to determine for themselves the efficacy of individual requests and the degree to which an underlying strategy may be technically sound, feasible, and defensible.

Requests Consistently Not Approved

In 2003-04, ED was consistent in not approving several categories of requested changes from states. The following list now might be considered the "non-negotiable" types of requests that have been consistently denied:

- The use results of out-of-level testing in AYP calculations, except if it is consistent with the 1 percent cap requirements. ED now considers out-of-level tests to be one form of alternate assessments that can be made available for use with severely cognitively disabled students consistent with the applicable regulations;
- Limiting identification for improvement to schools that missed AYP in the same subject and group for two consecutive years;
- Retroactive application of any amendments to years prior to the 2003-04 school year;
- Blanket exemptions from testing of students for medical-emergencies (decisions must be made on a case-by-case basis);
- Switching the order of choice and supplemental educational services—school choice must be offered in the first year of identification. Further, school choice cannot be limited to "non-proficient" students.

Unanticipated Approvals

Perhaps the most interesting decisions in the past year came in the form of the “Dear Colleague” letters to chief state school officers and other guidance documents released by ED. The new flexibility for the inclusion of LEP students, for example, was similar to requests some states had made in their original accountability plans—requests that had been denied, thus making their appearance in guidance rather unexpected.

Although the NCLB Act itself includes the provision to allow the use of results for only those students served or eligible to be served in schools that operate Targeted Assistance Title I programs, this practice had not received much attention until North Carolina specifically requested its use. It is not clear whether this will affect the fall 2004 AYP results for schools and districts there. Interestingly, ED limited exercise of this application to schools that had a minimum “n” of at least 40 but did not commit that to writing so the underlying rationale and policy are unknown.

To a lesser degree, but still somewhat surprising, was the approval Texas received to apply rounding rules to “all AYP elements.” Up to this point, some states—the number is unknown—may have been applying rounding rules but not to “count” elements such as participation rates. In this vein, however, it was a bit more surprising to learn that ED had signaled to at least one state its willingness to permit rounding up of any decimal to the next highest whole number.

Other unanticipated approvals included the following:

- Limiting identification of school districts for improvement such that they must miss AYP targets in each of three grade spans in a given content area for two years consecutively. (Thus, it appears that the other academic indicators need not be considered for district-level AYP decisions since these indicators are not the same at the elementary/middle and high school levels.);
- The use of confidence intervals for safe harbor calculations, albeit generally limited to 75 percent without a written rationale or stated policy from ED;
- Averaging data across “multiple years” in safe harbor determinations;
- Rounding rules in which decimals can be rounded up to the next whole number;
- The use of confidence intervals for participation and graduation rates;
- The use of a larger minimum “n” for large school districts (minimum “n” is the same as that used for schools until the district reaches a certain size, then increases proportionally with the size of the district);
- The use of longer than one calendar year to define full academic year for SWDs taking a state’s alternate assessments when the IEP indicates that a longer time period is needed to collect the achievement information;
- The inclusion of students no longer receiving special education services in the SWDs category as long as their IEPs call for continued monitoring; and
- The use of data for two consecutive years to make safe harbor determinations in small schools even though the minimum “n” requirements are not met.

Requests that Await Public Response

Many states have submitted amendment requests but, as of August 2004, have not yet received a formal response from ED with regard to one or more specific issues. These include the following:

- The use of results for a subset of grades until 2005-06 in states that already have assessments in grades 3-8 and want to make AYP determinations on the basis of student achievement results in only two of these grades;
- The use of results from “re-tests” that occur after official administrations in AYP calculations;
- The exclusion from calculation of participation rates students whose parents have formally withdrawn them from testing in states where this right is protected by law;

- The use of growth models in lieu of percent proficient or in safe harbor;
- The use of a “scaled” minimum “n” that would effectively exclude virtually all small schools in the state from standard AYP determinations substituting a quantitative and qualitative review process in its place; and
- The inclusion of students in only one program subgroup (low income, SWDs, or LEP).

Next Steps

For the most part, it appears that states only submitted requests for amendments in areas where they believed a possibility for approval existed, either due to previous patterns of approvals or to the preponderance of evidence or logic on their behalf. In some cases, states that were denied their requests may make adjustments in their rationale or gather additional evidence and resubmit their requests for amendment. Other states may accept the denials or may be waiting for changes in political receptiveness to amendment of the law itself—which will undoubtedly be the next chapter in this evolving story. In the meantime, some states will very probably submit additional amendments for ED review based on “surprise” decisions stemming from the current round of amendment approvals. When combined with the fall 2004 elections and on-going discussions of other federal legislation (e.g., reauthorization of the Individuals with Disabilities Education Act), the next year should prove to be another interesting one for those following federal education policy.

References

- Erpenbach, W. J., Forte Fast, E., and Potts, A. (2003). *Statewide educational accountability under NCLB: Central issues arising from an examination of state accountability workbooks and U.S. Department of Education reviews under the No Child Left Behind Act of 2001*. Washington, DC: The Council of Chief State School Officers. Retrieved September 21, 2004 from <http://www.ccsso.org/content/pdfs/StatewideEducationalAccountabilityUnderNCLB.pdf>.
- Final Regulations, Title I—Improving the Academic Achievement of the Disadvantaged; Final Rule [regarding the achievement of students with the most significant cognitive disabilities], *Federal Register*, 68698 (Dec. 9, 2003) (to be codified at 34 CFR Part 200).
- Final Regulations, Title I—Improving the Academic Achievement of the Disadvantaged; Standards and Assessments, *Federal Register*, (June 5, 2002) 34 CFR Part 200, pp. 45038-45047.
- Gewertz, C. (2003, October 8). Pa. adapts to new federal education law. *Education Week*, pp. 18, 20. Retrieved September 21, 2004 from <http://www.edweek.org/ew/ewstory.cfm?slug=06PA-AYP.h23> (registration required).
- Hill, R. and DePascale, C.A. (April 2003). *Adequate yearly progress under NCLB: Reliability considerations*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Chicago, IL. Retrieved September 21, 2004 from http://www.nciea.org/publications/NCMEconf_RHCD03.pdf.
- Linn, R. L., Baker, E. and Betebenner, D. (2002). *Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001*. CSE Technical Report 567. Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, University of California at Los Angeles. Retrieved September 21, 2004 from http://www.aera.net/pubs/er/pdf/vol31_06/AERA310602.pdf.
- Marion, S., White, C., Carlson, D., Erpenbach, W. J., Rabinowitz, S. and Sheinker, J. (2002). *Making valid and reliable decisions in determining adequate yearly progress*. Washington, DC: Council of Chief State School Officers. Retrieved September 21, 2004 from <http://www.ccsso.org/content/pdfs/AYPpaper.pdf>.
- The No Child Left Behind Act of 2001*. Public Law 107-110. Information available at <http://www.ed.gov/nclb/landing.jhtml>.
- Olson, L. (2004, July 14). Statesicker over changes to AYP plans. *Education Week*, p 1, 22. Retrieved September 21, 2004 from <http://www.edweek.org/ew/ewstory.cfm?slug=42Account.h23> (registration required).
- Paige, R. (2003, June 27). *Key policy letters signed by the education secretary or deputy secretary* [regarding inclusion of students with disabilities in AYP]. Washington, DC: U.S. Department of Education. Retrieved September 21, 2004 from <http://www.ed.gov/policy/speced/guid/secletter/030627.html>.
- Paige, R. (2004, February 20). *Key policy letters signed by the education secretary or deputy secretary* [regarding additional flexibility related to limited English proficient students]. Washington, DC: U.S. Department of Education. Retrieved September 21, 2004 from <http://www.ed.gov/policy/gen/guid/secletter/040220.html>.

- Paige, R. (2004, March 2). *Letter to chief state school officers regarding inclusion of students with disabilities in state accountability systems*. Washington, DC: U.S. Department of Education. Retrieved September 21, 2004 from <http://www.ed.gov/admins/lead/account/csso030204.html>.
- Paige, R. (2004, March 29). *Secretary Paige issues new policy for calculating participation rates under No Child Left Behind*. Washington, DC: U.S. Department of Education. Retrieved September 21, 2004 from <http://www.ed.gov/news/pressreleases/2004/03/03292004.html>.
- Paige, R. (2004, April 27). Letter to the honorable Edward M. Kennedy. Washington, DC: The Secretary of Education. Washington, DC: U.S. Department of Education. Retrieved September 21, 2004 from <http://www.ed.gov/policy/elsec/guid/secletter/040427.html>.
- Robelen, E. W. (2004, March 31). Schools seek participation on test days. *Education Week*, pp.1, 30. Retrieved September 21, 2004 from <http://www.edweek.org/ew/ewstory.cfm?slug=29Participation.h23> (registration required).
- Simon, R. J. (2004, January 26). *Letter to the honorable Susan Castillo, Oregon superintendent of public instruction*. Washington, DC: U.S. Department of Education. Retrieved September 21, 2004 from <http://www.ed.gov/admins/lead/account/letters/acor.html>.
- Simon, R. J. (2004, February 2). *Letter to the honorable Gene Wilhoit, Kentucky commissioner of education*. Washington, DC: U.S. Department of Education. Retrieved September 21, 2004 from <http://www.ed.gov/admins/lead/account/letters/acky.html>.
- Simon, R. J. (2004, February 5). *Letter to chief state school officers regarding April 1 for submittal of amendments*. Washington, DC: U.S. Department of Education. Retrieved September 21, 2004 from <http://www.ed.gov/admins/lead/account/amproc.doc>.
- Simon, R. J. (2004, March 4). *Letter to the honorable David Stewart, West Virginia superintendent of schools*. Washington, DC: U.S. Department of Education. Retrieved September 21, 2004 from <http://www.ed.gov/admins/lead/account/letters/acwv2.html>.
- Simon, R. J. (2004, May 20). *Letter to the honorable Michael E. Ward, North Carolina Department of Public Instruction*. Washington, DC: U.S. Department of Education. Retrieved September 21, 2004 from <http://www.ed.gov/admins/lead/account/letters/acnc.html>.
- Tomalis, R. J. (2003, December 31). *Letter to the honorable David Stewart, West Virginia superintendent of schools*. Washington, DC: U.S. Department of Education. Retrieved September 21, 2004 from <http://www.ed.gov/admins/lead/account/letters/acwv.html>.

Appendix A

State Accountability Workbook Decisions That Changed Between 2002 and 2004

Topic	Original Decision	Final Decision ¹	Comment
Dual accountability systems	Permitted only when a school could not achieve a high performance level if it was identified for improvement.	States can have dual systems that recognize school regardless of their AYP outcomes.	
Different minimum “n’s” for subgroups	Not permitted.	Permitted.	
Larger minimum “n’s” for school districts	Not permitted.	Models approved permitting the use of a minimum “n” which increases proportionally as the size of the district increases.	
Including scores for “exited” SWDs in subgroup AYP outcomes	Not permitted.	Permitted to the extent that monitoring services are required in student’s IEP.	
Including scores for “exited” LEP students in subgroup AYP outcomes	Not permitted.	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.	See Secretary’s February 20, 2004, policy letter on LEP students.
Use of scores for LEP students in AYP	Required.	Scores may be excluded from percent proficient calculations during an LEP student’s first year of enrollment in U.S. schools; if the student takes the English Language Proficiency test that year (as required), the student may be counted as a participant in the assessments.	See Secretary’s February 20, 2004, policy letter on LEP students.

¹ 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.

Topic	Original Decision	Final Decision¹	Comment
Out-of-level testing	Not permitted.	Permitted as an alternate assessment aligned to alternate standards and subject to 1 percent cap.	Modified in regulations adopted on December 9, 2003 (<i>Federal Register</i>). See also Secretary's March 2, 2004 policy letter on SWDs.
Alternate assessments for SWDs	Permitted but must be based on state's academic content and student academic achievement standards.	Alternate assessments based on alternate standards permitted, but not more than 1 percent of total student population can be counted as proficient in relation to the alternate standards.	Modified in regulations adopted on December 9, 2003 (<i>Federal Register</i>). See also Secretary's March 2, 2004 policy letter on SWDs.
Participation rate	Based on all students enrolled during a state's annual test window	States may average participation rates over a three-year period.	See Secretary's March 29, 2004, policy letter on participation rates.
Exemption from testing	Not permitted.	States permitted to exempt students from testing on a case-by-case bases due to significant medical emergencies.	See Secretary's March 29, 2004, policy letter on calculating participation rates.
Calculating percent proficient	Use number of students enrolled full academic year for the denominator.	Use number of students enrolled for a full academic year <u>and tested</u> for the denominator.	
Uniform averaging	Uniform averaging required.	States now 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.	
Use of statistical tests in safe harbor reviews	Not permitted.	Use of a 75% confidence interval allowed.	Generally limited to 75% in approvals after June 2003.
Safe harbor reviews for small schools	Unclear.	A state may conduct a review of current year and prior year performance even when its minimum "n" requirements are not met.	

Topic	Original Decision	Final Decision ¹	Comment
Use of statistical tests for participation and graduation rates	Not permitted for “count” elements.	Unclear.	States have been allowed to apply confidence intervals (CI) to AYP calculations. But when requests have been specifically for CIs around “count” items, they have been denied.
First test score when students have multiple opportunities to take an assessment.	First test score counts for AYP determinations.	States may now “bank” results when afforded multiple opportunities to test provided an “official” point has been designated at which students are expected to have attained the tested standards.	ED stated in comments with regulations adopted on December 9, 2003 (<i>Federal Register</i>) that states have more flexibility toward this end than originally understood.
Identification for improvement	School or district missing any AYP target two consecutive years is identified for improvement.	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.	
AYP decisions in school districts	Districts required to “roll up” data across schools for AYP determinations.	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.	
Full academic year extending beyond one calendar year for SWDs	Not permitted.	Full academic year for SWDs may extend beyond one year when IEP indicates that longer period is needed to collect achievement information.	
Graduation rate covering more than four years	Students taking more than four years to graduate could not be counted as graduates in calculating graduation rates.	SWDs who take more than four years to graduate, consistent with their IEPs may be counted as graduates.	
Rounding up in AYP decisions	Unclear.	States may round up when calculating AYP elements.	

Appendix B

Ten Principles for State Accountability Systems

The following are the ten principles of state accountability systems as specified in the June 2003 U.S. Department of Education-issued Consolidated State Application Accountability Workbook for State Grants under Title IX, Part C, Section 9302 of the Elementary and Secondary Education Act (Public Law 107-110).

Principle 1: All Schools

- 1.1 Accountability system includes all schools and districts in the state.
- 1.2 Accountability system holds all schools to the same criteria.
- 1.3 Accountability system incorporates the academic achievement standards.
- 1.4 Accountability system provides information in a timely manner.
- 1.5 Accountability system includes report cards.
- 1.6 Accountability system includes rewards and sanctions.

Principle 2: All Students

- 2.1 The accountability system includes all students.
- 2.2 The accountability system has a consistent definition of full academic year.
- 2.3 The accountability system properly includes mobile students.

Principle 3: Method of AYP Determinations

- 3.1 Accountability system expects all student subgroups, public schools, and LEAs to reach proficiency by 2013-14.
- 3.2 Accountability system has a method for determining whether student subgroups, public schools, and LEAs made adequate yearly progress.
 - 3.2a Accountability system establishes a starting point.
 - 3.2b Accountability system establishes statewide annual measurable objectives.
 - 3.2c Accountability system establishes intermediate goals.

Principle 4: Annual Decisions

- 4.1 The accountability system determines annually the progress of schools and districts.

Principle 5: Subgroup Accountability

- 5.1 The accountability system includes all the required student subgroups.
- 5.2 The accountability system holds schools and LEAs accountable for the progress of student subgroups.
- 5.3 The accountability system includes students with disabilities.
- 5.4 The accountability system includes limited English proficient students.
- 5.5 The State has determined the minimum number of students sufficient to yield statistically reliable information for each purpose for which disaggregated data are used.
- 5.6 The State has strategies to protect the privacy of individual students in reporting achievement results and in determining whether schools and LEAs are making adequate yearly progress on the basis of disaggregated subgroups.

Principle 6: Based on Academic Assessments

- 6.1 Accountability system is based primarily on academic assessments.

Principle 7: Additional Indicators

- 7.1 Accountability system includes graduation rate for high schools.
- 7.2 Accountability system includes an additional academic indicator for elementary and middle schools.
- 7.3 Additional indicators are valid and reliable.

Principle 8: Separate Decisions for Reading/Language Arts and Mathematics

- 8.1 Accountability system holds students, schools and districts separately accountable for reading/language arts and mathematics.

Principle 9: System Validity and Reliability

- 9.1 Accountability system produces reliable decisions.
- 9.2 Accountability system produces valid decisions.
- 9.3 State has a plan for addressing changes in assessment and student population.

Principle 10: Participation Rate

- 10.1 Accountability system has a means for calculating the rate of participation in the statewide assessment.

SCASS