



**Standards and Assessments
Peer Review Guidance:
Information and Examples for
Meeting Requirements of the
No Child Left Behind Act of 2001**



April 28, 2004

**U. S. Department of Education
Office of Elementary and Secondary Education
Washington, D.C. 20202**

TABLE OF CONTENTS

I. INTRODUCTION

STATUTORY AND REGULATORY REQUIREMENTS FOR NCLB ASSESSMENT SYSTEMS1

STATE ASSESSMENT SYSTEM DESIGN4

THE PEER REVIEW PROCESS5

ROLE OF PEER REVIEWERS.....6

REVIEW PROCESS6

REVIEW TEAMS.....7

STATE’S ROLE.....7

II. SECTION 1

A SINGLE STATEWIDE SYSTEM OF CHALLENGING ACADEMIC CONTENT STANDARDS APPLIED TO ALL PUBLIC SCHOOLS AND LEAS8

III. SECTION 2

A SINGLE STATEWIDE SYSTEM OF CHALLENGING ACADEMIC ACHIEVEMENT STANDARDS APPLIED TO ALL PUBLIC SCHOOLS AND LEAS14

IV. SECTION 3

A SINGLE STATEWIDE SYSTEM OF ANNUAL HIGH-QUALITY ASSESSMENTS.....23

V. SECTION 4

A SYSTEM OF ASSESSMENTS WITH HIGH TECHNICAL QUALITY.....32

VI. SECTION 5

ALIGNMENT OF ACADEMIC CONTENT STANDARDS, ACADEMIC ACHIEVEMENT STANDARDS, AND ASSESSMENTS41

Standards and Assessments Peer Review Guidance
April 28, 2004

VII. SECTION 6

INCLUSION OF ALL STUDENTS IN THE ASSESSMENT SYSTEM50

VIII. SECTION 7

AN EFFECTIVE SYSTEM OF ASSESSMENT REPORTS.....58

Standards and Assessments Peer Review Guidance
April 28, 2004

INTRODUCTION

Raising academic standards for all students and measuring student achievement to hold schools accountable for educational progress are central strategies for promoting educational excellence and equity in our schools. The *No Child Left Behind Act of 2001* (NCLB) reformed Federal educational programs to support State efforts to establish challenging standards, to develop aligned assessments, and to build accountability systems for districts and schools that are based on educational results. In particular, NCLB includes explicit requirements to ensure that students served by Title I are given the same opportunity achieve to high standards and are held to the same high expectations as all other students in each State.

Building on the foundation of standards and assessments required of States by the Improving America's Schools Act of 1994 (IASA), the current NCLB requirements include high-quality academic assessments, accountability systems, and teacher preparation and training aligned with challenging State academic standards so that students, teachers, parents, and administrators can measure progress against common expectations for students' academic achievement. NCLB extends assessment requirements to include, by school year 2005-06, annual assessments in reading/ language arts and mathematics in all grades 3 through 8 and assessments administered at least once in grades 10 through 12. In addition, States must develop academic content standards in science by 2005-06 and aligned assessments based on those standards by 2007-08. The science assessments must be administered at least once in each of three grade spans: 3-5, 6-9, and 10-12.

The purpose of this guidance is twofold: (1) to inform States about what would be useful evidence to demonstrate that they have met NCLB standards and assessments requirements; and (2) to guide teams of peer reviewers who will examine the evidence submitted by States and advise the Department as to whether a State has met the requirements. The intent is to help States develop comprehensive assessment systems that provide accurate and valid information for holding districts and schools accountable for student achievement against State standards. Although this document addresses each requirement separately, reviewers and States should recognize that the requirements are interrelated and that decisions about whether a State has met the requirements will be based on a comprehensive examination of the evidence submitted.

Statutory and Regulatory Requirements for NCLB State Assessment Systems

Under NCLB, States must develop challenging academic standards that have the following characteristics:

- Be the same academic standards that the State applies to all public schools and public school students in the State;

Standards and Assessments Peer Review Guidance
April 28, 2004

- Include the same knowledge, skills, and levels of achievement expected of all students; and
- Include at least mathematics, reading/language arts, and, beginning in the 2005-2006 school year, science.

Academic **content** standards must specify what all students are expected to know and be able to do; contain coherent and rigorous content; and encourage the teaching of advanced skills. A State's academic content standards may either be grade-specific or may cover more than one grade if grade-level content expectations are provided for each of grades 3 through 8. At the high school level, the academic content standards must define the knowledge and skills that all high school students are expected to have in at least reading/language arts, mathematics, and, beginning in the 2005-06 school year, science, irrespective of course titles or years completed.

Academic **achievement** standards must be aligned with the State's academic content standards. For each content area, a State's academic achievement standards must include at least two levels of achievement (proficient and advanced) that reflect mastery of the material in the State's academic content standards, and a third level of achievement (basic) to provide information about the progress of lower-achieving students toward mastering the proficient and advanced levels of achievement.

For each achievement level, a State must provide descriptions of the competencies associated with that achievement level and must determine the assessment scores ("cut scores") that differentiate among the achievement levels. The State must also provide a description of the rationale and procedures used to determine each achievement level. Unlike content standards, which may address a cluster of grade levels, academic achievement standards must be developed for each grade and subject assessed, even if the State's academic content standards cover more than one grade.

With respect to academic achievement standards in science, a State must develop achievement levels and descriptions no later than the 2005-06 school year and must determine "cut scores" after the State has developed its science assessments, but no later than the 2007-08 school year.

Under NCLB, the State assessment system must have the following characteristics:

- Assessments must be aligned with State academic content and achievement standards, and they must provide coherent information about student attainment of State standards in at least mathematics and reading/language arts. Beginning in 2007-08, the system must also include assessments in science.
- The same assessment system must be used to measure the achievement of all students.

Standards and Assessments Peer Review Guidance
April 28, 2004

- The assessment system must be designed to be valid and accessible for use by the widest possible range of students, including students with disabilities and students with limited English proficiency (LEP).
- Initially, assessments must be administered annually to students in at least one grade in each of three grade ranges--grades 3 through 5, grades 6 through 9, and grades 10 through 12. Beginning in 2005-06, the mathematics and reading/language arts assessments must be given in each of grades 3 through 8 in addition to one of the grades 10 through 12.
- The assessment system must provide for--
 - Participation of all students in the grades being assessed;
 - Reasonable adaptations and appropriate accommodations for students with diverse learning needs, where such adaptations or accommodations are necessary to measure the achievement of those students relative to State standards; and
 - Inclusion of LEP students, who must be assessed in a valid and reliable manner and provided reasonable accommodations including, to the extent practicable, assessments in the language and form most likely to yield accurate and reliable information on what they know and can do in academic content areas, until such students have achieved English language proficiency; except that the reading/language arts achievement of any student who has attended school in the United States for three consecutive years must be tested in English.
- The assessment system must involve multiple approaches with up-to-date measures of student achievement, including measures that assess higher-order thinking skills and understanding of challenging content.
- Assessments must be valid and reliable for the purposes for which the assessment system is used and be consistent with relevant, nationally recognized professional and technical standards.
- The assessment system must be supported by evidence from test publishers or other relevant sources that the assessment system is of adequate technical quality for each purpose required under the Act.
- The assessment system must objectively measure academic achievement, knowledge, and skills without evaluating or assessing personal or family beliefs and attitudes, except that this provision does not preclude the use of constructed-response, short answer, or essay questions, or items that require a student to analyze a passage of text or to express opinions.
- Assessment results must be disaggregated within each school and district by gender, major racial and ethnic groups, English proficiency status, migrant status, students with disabilities as compared to students without disabilities, and economically disadvantaged students as compared to students who are not economically disadvantaged. Such disaggregation is not required when the number of students in a category is insufficient to yield statistically reliable information or if the results would reveal personally identifiable information about an individual student.

Standards and Assessments Peer Review Guidance

April 28, 2004

- The assessment system must provide individual student interpretive, descriptive, and diagnostic reports that include individual scores or other information on the attainment of student achievement standards and help parents, teachers, and principals to understand and address the specific academic needs of students. These reports must be provided as soon as practicable after the assessment is given and in an understandable and uniform format.

Under NCLB, the statewide assessment system will be the primary means for determining whether schools and school districts are making adequate yearly progress (AYP) toward educating students to high standards. In determining the progress of schools, States must include scores of all students enrolled in the school for at least a full academic year. In determining the progress of school districts, States must include scores of all students enrolled in schools in the district for a full academic year, even if they have attended several different schools.

Because NCLB makes the State assessment system central to holding schools and districts accountable, this document focuses on the uses of the State assessment system at the school and district levels. Nevertheless, peer reviewers should note that the State assessment system is also required to report results at the level of individual students.

State Assessment System Design

A State may include in its academic assessment system either (or both) criterion-referenced assessments and assessments that yield national norms, provided that, if the State uses only assessments referenced against national norms at a particular grade, those assessments are augmented with additional items as necessary to measure accurately the depth and breadth of the State's student academic achievement standards.

A State that includes a combination of criterion and norm-referenced assessments in its assessment system must demonstrate that the system has a rational and coherent design that:

- Identifies the assessments to be used;
- Indicates the relative contribution of each assessment towards ensuring alignment with the State's academic content standards and toward determining the adequate yearly progress of each school and local educational agency (LEA); and
- Provides information regarding the progress of students relative to the State's academic standards.

A State's assessment system may employ either a uniform set of assessments statewide or a combination of State and local assessments. States using a combination of State and local tests must address issues of comparability and equivalency. For example, will proficiency on one local assessment be comparable to proficiency on another local assessment? Additionally, States must consider how they will aggregate to the State level the results from local assessments, as is required by NCLB.

Standards and Assessments Peer Review Guidance

April 28, 2004

States that choose to include a combination of State and local assessments will need to demonstrate that their system has a rational and coherent design that--

- Identifies the assessments to be used at the State and local levels;
- Indicates the relative contribution of each assessment toward ensuring alignment with the State's academic content standards and toward determining the adequate yearly progress of each school and LEA; and
- Provides information regarding the progress of students relative to the State's academic standards.

Further, a State that includes local assessments must also--

- Establish technical criteria to ensure that each local assessment addresses the depth and breadth of the State's academic standards; is valid, reliable, and of high technical quality; expresses student results in terms of the State's academic achievement standards; and is designed to provide a coherent system across grades and subjects.
- Demonstrate that all local assessments are equivalent in their content coverage, difficulty, and quality to one another and to State assessments; have comparable validity and reliability with respect to groups of students described in section 1111(b)(2)(C)(v); and provide unbiased, rational, and consistent determinations of the annual progress of schools and LEAs within the State.
- Review and approve each local assessment to ensure that it meets or exceeds the State's technical quality for assessments.
- Be able to aggregate, with confidence, data from local assessments to determine whether the State has made adequate yearly progress.

In implementing their assessment system, States have two main responsibilities: (1) they must develop, score, and report findings from State assessments, and (2) they must promulgate rules and procedures for local assessment systems if the State has such systems, as well as monitor them, to ensure technical quality and compliance with Title I requirements. The second function is particularly significant in assessment systems with strong local responsibility.

The Peer Review Process

To determine whether States have met NCLB standards and assessments requirements, the U.S. Department of Education (ED) will use a peer review process involving experts in the fields of standards and assessments. The review will evaluate States' assessment systems only against NCLB requirements. In other words, reviewers will examine characteristics of a State's assessment system that will be used to hold schools and school districts accountable under NCLB. They will not assess compliance of States' assessment systems with other Federal laws such as Title VI of the Civil Rights Act of 1964, Section 504 of the Rehabilitation Act of 1973, or provisions of the Individuals with Disabilities Education Act (IDEA). The fact that an assessment system meets NCLB assessment requirements does not necessarily mean that it complies with other laws. For guidance on compliance with Federal civil rights laws, States may consult with the Department of

Standards and Assessments Peer Review Guidance

April 28, 2004

Education's Office for Civil Rights. For guidance on compliance with the IDEA, States may consult with the Office of Special Education Programs.

Furthermore, the peer review process will not directly examine a State's academic standards, assessment instruments, or specific test items. Rather, it will examine *evidence* compiled and submitted by each State that is intended to show that its assessment system meets NCLB requirements. Such evidence may include, but is not limited to, results from alignment studies; results from validation studies; written policies if appropriate, on providing accommodations for students with disabilities and LEP students; written policies on native-language testing of LEP students (if applicable); and score reports showing disaggregation of student achievement data by the statutorily specified student subgroups. Peer reviewers will advise the Department on whether a State assessment system meets a particular requirement based on the totality of evidence submitted. Peer reviewers will also provide constructive feedback to help States strengthen their assessment systems.

Role of Peer Reviewers

Using this *Guidance* as a framework, the peer reviewers will provide their expert professional judgment, based on evidence supplied by the State, of the degree to which the State's final assessment system complies with the requirements of Title I. Their evaluation of the final assessment system will serve two purposes. First, the peer reviewers' consensus comments will be sent to the State as a technical assistance tool to support improvements in the system. Second, the peer reviewers' comments will inform the decision of the Assistant Secretary for Elementary and Secondary Education regarding approval of each State assessment system.

Review Process

The materials submitted to the Department by the State are sent to each member of the peer review team in advance of a review meeting to allow for a thorough independent review based on the *Guidance*. At the review meeting, the team of at least three peer reviewers discusses a State's system, as represented by the evidence provided by the State, and records a consensus opinion.

For the Final Assessment Review, evidence means documents such as actual statutes, State regulations, test administration manuals, board resolutions, or assessment reports. Sufficient evidence must be provided to convince an experienced professional that the assessment system is being implemented in a manner that meets NCLB requirements.

This *Guidance* is a framework used to make a series of analytic judgments. Reviewers will address each of the peer reviewer questions in the *Guidance*, evaluating the status of each component of the system on the basis of the documentation provided by the State. A brief statement of the degree to which the assessment system meets the NCLB requirements and the changes needed, if any, summarizes this analytic examination of the assessment system.

Standards and Assessments Peer Review Guidance

April 28, 2004

Review Teams

The peer review team prepares a consensus report based on its examination of the materials submitted by the State. In each team, one person will be designated team leader; this person is responsible for seeing that consensus notes are clear, complete, and delivered to ED staff at the end of the review meeting. The peer reviewers are responsible for providing feedback to each State that is informative and is consistent with professional standards and best practice. Generally, if changes in a State assessment system are required in order to meet Title I requirements, peer reviewers will present options rather than prescriptive instructions.

An ED staff person, assigned as a resource to each team, is responsible for assisting the review team in obtaining adequate and appropriate information from the State prior to the review meeting; contacting the State during the review meeting to obtain clarification or additional information needed by the reviewers; securing resources needed to support the team during the meeting; and accurately reporting the review team's deliberations as ED determines the State's compliance status. ED staff may question, or even challenge, the peer reviewers in order to promote clarity and consistency with the *Guidance*; they will not, however, impose their views or require substantive changes in the peer reviewers' consensus report.

States are invited to submit evidence of NCLB compliance consistent with the peer review schedule to be announced by the Department.

State's Role

To facilitate the peer review process, a State should organize its evidence with a brief narrative response to each of the "peer reviewer questions" in the *Guidance*. The Department will provide a template to the State to help organize supporting documents that constitute evidence of meeting the assessment requirements. The State will be asked to designate staff who can be contacted by phone during the review to provide clarification.

Standards and Assessments Peer Review Guidance
April 28, 2004

Section 1: A single statewide system of challenging academic content standards applied to all public schools and LEAs.

Reference in NCLB legislation: Sec. 1111(b)(1)
Reference in final regulations: Sec. 200.1

Overview

As the starting point for establishing a high quality assessment and accountability system under NCLB, States must develop a set of challenging academic content standards that define what all public school students in the State are expected to know and be able to do. A State’s academic content standards are to be applied to all public elementary and secondary school students.

The table below provides a summary of the content, grade level, and timeline requirements for the academic content standards.

Content Area	Grade levels	Due	Notes
Reading/language arts	<ul style="list-style-type: none"> • Each grade: 3 - 8; <li style="text-align: center;">and • Grade range: 10 - 12 	May 2003	<ul style="list-style-type: none"> • If a State’s standards cover grade ranges (e.g., 3 - 5 and 6 - 8) rather than the specific grades, 3 - 8, the State must develop grade-specific expectations in addition to its standards. • At the high school level, standards must define the knowledge and skills that are expected of all students prior to graduation. They may be linked to specific courses if all students must take these courses in order to graduate.
Mathematics	<ul style="list-style-type: none"> • Each grade: 3-8; <li style="text-align: center;">and • Grade range: 10-12 		
Science	<ul style="list-style-type: none"> • Grade ranges: 3 - 5; 6 - 9; 10 - 12 	By the 2005 - 2006 school year	<ul style="list-style-type: none"> • At the high school level, standards must define the knowledge and skills that are expected of all students prior to graduation. They may be linked to specific courses if all students must take these courses in order to graduate.

These standards must be rigorous and encourage the teaching of advanced skills. This means that a State should not adopt “minimum competency” standards or otherwise encourage low expectations for any students. Further, these standards must be coherent. That is, they must include only content that is meaningful with regard to the “domain”, that is appropriate for the grade level specified, and that reflects clearly articulated progressions across grade levels.

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 1: CONTENT STANDARDS

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>1.1</p> <p>(a) Has the State formally approved/adopted, by May 2003, challenging academic content standards in reading/language arts and mathematics that –</p> <ul style="list-style-type: none"> • cover each of grades 3-8 and the 10-12 grade range, <i>or</i> • if the academic content standards relate to grade ranges, include specific content expectations for each grade level? <p>AND</p> <p>(b) Are these academic content standards applied to <i>all</i> public schools and students in the State?</p>	<p>The State has formally approved/adopted academic content standards for all students in reading/ language arts and mathematics that are specific to each grade level 3 through 8 and for the 10-12 grade range, and represent the full range of knowledge and skills that students should be expected to know and be able to do by the time they graduate.</p> <p>The State has formally approved/adopted academic content standards or frameworks in reading/language arts and mathematics for the 3-5, 6-9, and 10-12 grade ranges. These standards or frameworks include grade-specific content expectations for all students in each grade level between 3 and 8 and for specific reading/language arts and mathematics courses, or combinations of courses that all students must take in the 10-12 grade range.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Written documentation in the form of State Board of Education minutes, regulations, official reports, letters or memoranda from the State to the LEAs, or other existing documents (i.e., not written by the State only to fulfill the requirements of the peer review process). Or, in States where the Chief State School Officer has the power to approve standards, written documentation of the formal approval of the final form of the reading/language arts and mathematics standards. • Letter from ED approving the content standards for grades 3 through 8 and high school if these content standards have not 	<p>The State has developed academic content standards but these standards have not been formally approved/adopted by the State.</p> <p>The State has formally approved/adopted academic content standards in reading/language arts but not in mathematics.</p> <p>The State has formally approved/adopted academic content standards in both reading/language arts and mathematics but these standards do not include grade-specific content expectations.</p> <p>At the high school level, the State’s formally approved/adopted standards provide only course descriptions for courses that some, but not all, students take in the 10-12 grade range.</p> <p>These descriptions do not represent the full range of knowledge and skills that students should be expected to know and be able to do by the time they graduate.</p> <p>The State does not clearly state how its academic content standards are to be applied to charter schools and other special purpose schools, such as detention centers, residential centers, and schools that serve students with special needs such as students with disabilities or students with limited English proficiency.</p> <p>The State’s statutes, policies, and guidance documents do not specifically state that its academic content standards apply to all public school students, including students with disabilities and students who are not proficient in</p>

Standards and Assessments Peer Review Guidance
April 28, 2004

	<p>been changed in any way since approval.</p> <ul style="list-style-type: none">• Documents that include or are based on the academic content standards explicitly address the needs of students with disabilities and limited English proficient students.	<p>English.</p>
--	--	-----------------

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 1: CONTENT STANDARDS

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>1.2 Has the State formally approved/adopted, academic content standards in science for elementary (grades 3-5), middle (grades 6-9), and high school (grades 10-12)? This must be completed by school year 2005-2006.</p>	<p>The State has formally approved/adopted academic content standards in science for the 3-5, 6-9, and 10-12 grade ranges.</p> <p>The State has formally approved/adopted academic content standards or frameworks in science for the 3-5 and 6-9 grade ranges and for a specific science course or combination of courses that all students must take in the 10-12 grade range in order to graduate from high school.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Written documentation in the form of State Board of Education minutes, regulations, official reports, letters or memoranda from the State to the LEAs, or other existing documents (i.e., not written by the State only to fulfill the requirements of the peer review process). Or, in States where the Chief State School Officer has the power to approve standards, written documentation of the formal approval of the final form of the science standards. 	<p>The State has developed academic content standards in science but these standards have not been formally approved/adopted by the State.</p> <p>The State has defined a process and timeline for developing and formally approving/adopting academic content standards in science but has not completed this process.</p> <p>At the high school level, the State’s formally approved/adopted standards provide only course descriptions that do not represent the full range of knowledge and skills that students should be expected to know and be able to do by the time they graduate.</p> <p>At the high school level, the State’s formally approved/adopted standards provide only course descriptions for courses that are not required for high school graduation.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 1: CONTENT STANDARDS

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>1.3 Are these academic content standards challenging? Do they contain coherent and rigorous content and encourage the teaching of advanced skills?</p>	<p>The State has a process for the development of academic content standards that includes expectations for higher grade levels that build upon and extend beyond the expectations for lower grade levels and incorporate higher-order thinking skills and understanding.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • A detailed description of the process the State used in developing its standards to review their rigor, such as its participation in a process to benchmark them to nationally recognized standards. This process should include substantive input from relevant stakeholders and individuals or organizations with expertise in standards development. 	<p>The State has developed a process and begun a plan for the development of academic content standards, but has not completed the process.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 1: CONTENT STANDARDS

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>1.4 Did the State involve education stakeholders in the development of its academic content standards?</p>	<p>The State’s process for developing its academic content standards involved diverse panels of educators, higher education representatives, parents, and community members familiar with the instructional needs of students with disabilities and students with limited English proficiency, as well as public hearings and consideration of public commentary on the standards.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Descriptions of the composition of groups involved in the development of the academic content standards indicate that, relative to all stakeholders in the population in general, a broad range of stakeholders was represented in the development process. • A description of how the standards were developed with input from many people in and outside of education. Such input might come through committees of curriculum, instruction, and content specialists <i>and</i> also from public hearings, public comment, or public review. • An assurance of sufficient diversity in the composition of groups involved in the development of the standards, including individuals knowledgeable of and concerned about the various categories of students with special needs, such as students with disabilities or students with limited English proficiency. 	<p>The State’s process for developing its academic content standards involved only K-12 educators and staff from the State Department of Education.</p> <p>The make-up of the State’s academic content standards’ panels did not include representatives of students with special needs, such as students with disabilities or students with limited English proficiency, or otherwise reflect the diversity of the State’s population.</p> <p>The State’s process for developing its academic content standards did not include an opportunity for public review and feedback.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

Section 2: A single statewide system of challenging academic achievement standards applied to all public schools and LEAs.

Reference in NCLB legislation: Sec. 1111(b)(1)
Reference in final regulations: Sec. 200.1

Overview

To establish the level of achievement a State expects of all public schools and LEAs, the NCLB requires States to develop a set of challenging academic achievement standards for every grade and content area assessed. These standards are to be applied to all public schools and LEAs and ensure inclusion of those students with disabilities and students who are not yet proficient in English.

Achievement Levels

Academic achievement standards for each grade-and-content area combination must include at least three achievement levels, which the State may label ‘proficient,’ ‘advanced,’ and ‘basic.’ Of these levels, proficient and advanced must represent high achievement and basic must represent achievement that is not yet proficient. These labels may vary from State to State, such as “meeting and mastering” the State standards that would equate to the proficient and advanced labels as described in the statute. A State may use more than three levels, but must clearly indicate which level represents the proficient performance expected of all students.

Descriptors and Cut Scores

In addition to these levels, the State’s academic achievement standards must include descriptions of the content-based competencies associated with each level. The State must also determine which specific scores on its assessments distinguish one level from another. These “cut scores” must be established through a process that involves both expert judgments and consideration of assessment results.

Alignment

As a set, the academic achievement standards must be aligned with the State’s academic content standards in that they capture the full range and depth of knowledge and skills defined in the State’s challenging, coherent, and rigorous academic content standards.

Timeline

Academic achievement standards in reading/language arts and mathematics for each of grades 3 through 8 and the 10-12 grade range must be in place by the 2005-06 school year. Academic achievement descriptors for science in grade spans 3-5, 6-9, and 10-12 must be in place by the 2005-06 school year and cut scores for science by the 2007-08 school year. States can develop the level and description components of the standards prior to the availability of assessment data that will be necessary to set the cut score components of these standards.

Standards and Assessments Peer Review Guidance April 28, 2004

Alternate academic achievement standards

A State is permitted to define alternate achievement standards to evaluate the achievement of students with the most significant cognitive disabilities and to give equal weight to a limited number of “proficient” assessment results based on alternate achievement standards in calculating adequate yearly progress (AYP). Alternate achievement standards must be aligned with the State’s academic content standards (i.e., include knowledge and skills that link to grade-level expectations), must promote access to the general curriculum, and must reflect professional judgment of the highest learning standards possible for the group of students with the most significant cognitive disabilities. The State defines alternate achievement standards through a documented and validated standards-setting process similar to the process used to establish achievement standards on the regular assessments.

As a State expands the regular assessments to include grades 3 through 8, it must also provide alternate assessments at grades 3 through 8. If these alternate assessments are based on grade-level achievement standards, they will include the same grade-level content as the test for which they are an alternate. The assessment procedures may differ from the regular assessment (e.g., body of work or performance tasks instead of multiple choice) but proficiency on these alternates is comparable to proficient performance on the regular assessment for the same grade. The State must provide evidence of comparability and be able to aggregate the results with results from the regular assessment.

For alternate assessments in grades 3 through 8 based on alternate achievement standards, the assessment materials should show a clear link to the content standards for the grade in which the student is enrolled although the grade-level content may be reduced in complexity or modified to reflect pre-requisite skills. For each grade, the State may define one or more alternate achievement standards for proficiency.

For students with the most significant cognitive disabilities who are mainstreamed, the concept of alternate achievement standards related to grade level may be ambiguous. For practitioners, the question is whether the alternate achievement standards for this group of students must be clearly different from grade to grade. The alternate achievement standards should be defined in a way that supports individual growth because of their linkage to different content across grades. When examined across grades, however, the alternate achievement standards are not likely to show the same clearly defined advances in cognitive complexity as the achievement standards set for the regular test or an alternate assessment based on grade-level standards. States are expected to rely on the judgment of experienced special educators and administrators, higher education representatives, and parents of students with disabilities as they define alternate achievement standards and to define alternate achievement standards in a manner that provides an appropriate challenge for students with the most significant cognitive disabilities as they move through their schooling.

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 2: ACADEMIC ACHIEVEMENT STANDARDS

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>2.1 Has the State formally approved/adopted challenging academic achievement standards in reading/language arts and mathematics for each of grades 3 through 8 and for the 10-12 grade range? These must be completed by school year 2005-2006.</p> <p>Has the State also approved/adopted alternate academic achievement standards for students with the most significant cognitive disabilities in reading/language arts and mathematics for each of the grades 3 through 8 and for the 10-12 grade span? If alternate achievement standards have not been developed and approved, then the alternate assessments for students with disabilities must be based on grade-level achievement standards.</p> <p>For students under section 602(3) of the Individuals with Disabilities Education Act with the most significant cognitive disabilities who take an alternate assessment, a State may, through a documented and validated standards-setting process, define alternate academic achievement standards, provided those standards (1) are aligned with the State’s academic content standards; (2) promote access to the general curriculum; and (3) reflect professional judgment of the highest achievement standards possible.</p>	<p>The State has formally approved/adopted academic achievement standards in reading/language arts and mathematics that are specific to each grade level between 3 through 8 and for the 10-12 grade span.</p> <p>The State has formally approved/adopted alternate achievement standards in reading/language arts and mathematics for the tested grades 3 through 8 and for the 10-12 grade span.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Written documentation in the form of State Board of Education minutes, regulations, official reports, letters or memoranda from the State to the LEAs, or other existing documents (i.e., not written by the State only to fulfill the requirements of the peer review process). Or, in States where the Chief State School Officer has the power to approve standards, written documentation of the formal approval of the final form of the reading/language arts and mathematics standards. • Documentation that illustrates how alternate achievement standards were defined and explains how they are linked to grade-level content. • A description of the process for revising cut scores and subsequent academic achievement standards <i>and</i> documentation that these revisions have been approved or formally submitted for approval by ED. 	<p>The State has developed academic achievement standards in reading/language arts and mathematics that are specific to each grade level between 3 and 8 and for the 10-12 grade range, but these standards have not been formally approved/adopted by the State.</p> <p>The State has formally approved/adopted academic achievement standards in reading/language arts but not in mathematics.</p> <p>The State has not approved/adopted grade-specific academic achievement standards in reading/language arts and mathematics.</p> <p>The State has developed alternate achievement standards, but the alternate achievement standards are not linked to State content standards in reading/language arts and mathematics.</p>

List State Evidence Here

**Standards and Assessments Peer Review Guidance
April 28, 2004**

SECTION 2: ACADEMIC ACHIEVEMENT STANDARDS

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>2.2 Has the State formally approved/adopted academic achievement descriptors in science for each of the grade spans 3-5, 6-9, and 10-12 (not due until school year 2005-2006)?</p> <p>Has the State formally approved/adopted academic achievement standards cut scores in science for each of the grade spans 3-5, 6-9, and 10-12 (not due until school year 2007-08)?</p> <p>Has the State also approved alternate academic achievement standards for students with the most significant cognitive disabilities in science for each of the grade spans tested? If alternate achievement standards have not been developed and approved, then the alternate assessments for students with disabilities must be based on grade-level achievement standards.</p>	<p>The State has formally approved/adopted academic achievement standards in science for each of the grade spans 3-5, 6-9, and 10-12.</p> <p>The State has formally approved/adopted alternate academic achievement standards in science for students with the most significant cognitive disabilities.</p> <p>Possible Evidence</p> <p>Written documentation in the form of State Board of Education minutes, regulations, official reports, letters or memoranda from the State to the LEAs, or other existing documents (i.e., not written by the State only to fulfill the requirements of the peer review process). Or, in States where the Chief State School Officer has the power to approve standards, written documentation of the formal approval of the final form of the science standards.</p>	<p>The State has developed academic achievement standards in science for each of the grade spans 3-5, 6-9, and 10-12 but these standards have not been formally approved/adopted by the State.</p> <p>The State has developed alternate achievement standards, but the alternate achievement standards are not linked to State content standards in science.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 2: ACADEMIC ACHIEVEMENT STANDARDS

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>2.3</p> <p>1. Do these academic achievement standards (and alternate achievement standards, if applicable) include for each content area –</p> <p>(a) At least three levels of achievement, including two levels of high achievement (proficient and advanced) that determine how well students are mastering a State’s academic content standards and a third level of achievement (basic) to provide information about the progress of lower-achieving students toward mastering the proficient and advanced levels of achievement; <u>and</u></p> <p>(b) descriptions of the competencies associated with each achievement level; <u>and</u></p> <p>(c) assessment scores (“cut scores”) that differentiate among the achievement levels and a rationale and procedure used to determine each achievement level?</p> <p>2. And, if the State has adopted alternate academic achievement standards, has it documented that it has--</p> <p>(a) implemented guidelines for IEP teams to apply in deciding when an individual student should be assessed on the basis of alternate achievement standards;</p> <p>(b) ensured that parents are informed when a child’s achievement will be based on alternate achievement standards and any possible consequences imposed by the LEA or State;</p> <p>(c) reported separately the number and percent of those students with disabilities assessed against alternate achievement standards, those assessed on an alternate assessment against grade-level standards, and those included in the regular assessment (including those administered with appropriate accommodations);</p> <p>(d) documented that students with the most significant cognitive disabilities are included in the general curriculum to the extent possible;</p> <p>(e) taken steps to promote use of appropriate accommodations for students tested against grade-level standards; and</p> <p>(f) provided information for teachers and other staff</p>	<p>The State has formally approved/adopted academic achievement standards that comprise three (or more) levels of achievement, each of which is associated with a description of the competencies expected of each required grade or grade range in high school and delineated by specific scores on the aligned assessment. Rationale and procedures for setting cut scores includes the consideration of impact data but is based primarily on expert judgments about content-based expectations. (<i>See elements 2.1 and 2.2 for required grades and grade ranges.</i>)</p> <p>The State has academic achievement standards (including alternate achievement standards for States that elect to use them) for every grade and subject assessed, even if the State’s academic content standards cover more than one grade.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Documentation in the form of State Board of Education minutes, official reports, letters or memoranda to the LEAs, (i.e., not written by the State only to fulfill the requirements of the peer review process) that refers to the levels, descriptions, and cut scores that make up the State’s academic achievement standards, and alternate achievement standards (if applicable). • Documentation that students with the most significant cognitive disabilities are included in the general curriculum to the extent possible. 	<p>The State has developed academic achievement standards that comprise three (or more) levels of achievement but these standards have not been formally approved/adopted.</p> <p>The State has formally approved/adopted academic achievement standards that comprise only two levels of achievement.</p> <p>The State’s academic achievement standards do not reflect the full range of the content expectations for each grade level, as specified in the State’s academic content standards.</p> <p>Rationale and procedures for setting cut scores were based solely or substantially on impact data rather than on content-based expectations.</p> <p>The State has used national norms as the sole basis for setting cut scores.</p> <p>The State has relied on “instructional level” as the basis for interpreting scores.</p>

Standards and Assessments Peer Review Guidance
April 28, 2004

regarding appropriate test administration practices, including use of accommodations for students with the most significant cognitive disabilities?	<ul style="list-style-type: none">• Training materials or guidelines for IEP teams to apply when deciding whether an individual student should be assessed on the basic of alternate achievement standards.	
---	---	--

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 2: ACADEMIC ACHIEVEMENT STANDARDS

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>2.4 Are the academic achievement standards (including alternate achievement standards as applicable) applied to <i>all</i> public elementary and secondary schools and <i>all</i> public school students in the State?*</p>	<p>The State’s statutes, policies, and guidance documents specifically state that its academic achievement standards apply to all public elementary and secondary school students, including students with disabilities and students who are not proficient in English.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Written documentation in the form of State Board of Education minutes or letters/memoranda from the Chief State School Officer to local superintendents stating the policy that the State’s academic achievement standards apply to all students. • Documentation and prefacing text for the standards use the word “all” consistently and inclusively and reflect other inclusive terminology. 	<p>The State’s statutes, policy, and guidance documents do not specifically state that its academic achievement standards apply to all public elementary and secondary school students, including students with disabilities and students who are not proficient in English.</p>

**OSEP guidance and NCLB requirements indicate that a student placed in a private school by a public agency for the purpose of receiving special education services must be included in the state assessment and their results attributed to the public school or LEA responsible for the placement.

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 2: ACADEMIC ACHIEVEMENT STANDARDS

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>2.5 How has the State ensured alignment between challenging academic content standards and the academic achievement standards?</p> <p>If the State has adopted alternate achievement standards, how has the State ensured alignment between its academic content standards and the alternate academic achievement standards?</p>	<p>The State’s academic achievement standards fully reflect its academic content standards for each required grade and describe what content-based expectations each achievement level represents. The ‘proficient’ achievement level represents attainment of grade-level expectations for that academic content area. The descriptors clearly define the skills for the attainment of that level.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • A description of the process used to develop each component (levels, descriptions, and cut scores) of the academic achievement standards that indicates that these standards were developed specifically to reflect the knowledge and skills in the academic content standards for each grade and content area combination. For alternate achievement standards, the State demonstrates that they reflect knowledge and skills that could lead to the content appropriate for the grade in which a student is enrolled. • Written documentation designed to accompany or explain the standards that delineate the content-based relationships between the academic achievement standards, alternate achievement standards, if any, and the academic content standards. 	<p>The State’s academic achievement standards do not fully reflect its academic content standards for each required grade nor describe what content-based expectations each achievement level represents.</p> <p>The achievement level that represents ‘Proficient’ defines performance that does not represent grade-level attainment of the content standards.</p> <p>Alternate achievement standards are restricted to functional life skills and are not linked in a meaningful way to academic knowledge/skills specified in State standards.</p>

List State Evidence

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 2: ACADEMIC ACHIEVEMENT STANDARDS

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>2.6 How did the State document involvement of diverse stakeholders in the development of its academic achievement standards and (if applicable) its alternate achievement standards?</p>	<p>The State’s process for developing its academic achievement standards, including alternate achievement standards, if used, involved diverse panels of educators, higher education representatives, parents, and community members, as well as public hearings and consideration of public commentary on the standards, if appropriate.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Descriptions of the composition of groups involved in the development of the academic achievement standards indicate that, relative to all stakeholders in the population in general, a broad range of stakeholders was represented in the development process. For development of its alternate achievement standards, experienced special education teachers and administrators, regular classroom teachers and parents of students with disabilities were included. • A description of how the standards, including alternative achievement standards, if used, were developed with input from many people in and outside of education. Such input might come through committees of curriculum, instruction, and content specialists <i>and</i> also from public hearings, public comment, or public review. • Evidence of diversity in the composition of groups involved in the development of the standards, including individuals knowledgeable of and concerned about the various categories of special needs students, e.g. students with disabilities and limited English proficient students. 	<p>The State’s process for developing its academic achievement standards involved only K-12 educators and staff from the State Department of Education.</p> <p>The make-up of the State’s academic achievement standards panels did not include representatives of students with special needs, such as students with disabilities or students with limited English proficiency or otherwise reflect the diversity of the State’s population.</p> <p>The State’s process for developing its academic achievement standards did not include an opportunity for public review and feedback.</p>

List State Evidence

Standards and Assessments Peer Review Guidance
April 28, 2004

Section 3: A single statewide system of annual high-quality assessments

Reference in NCLB legislation:	Sec. 1111(b)(3)
Reference in final regulations:	Sec. 200.2, 200.3, 200.5

Overview

To ensure that States are able to evaluate whether all students are achieving to high levels, NCLB requires States to develop a single statewide system of high quality assessments. All public school students must participate in this assessment system, including those with disabilities and those who are not yet proficient in English, so States must make their assessment system fully accessible to all students, (see Principle 6 for more information about inclusion). States must employ the same assessment system for all their public elementary and secondary schools and students.

States must have the reading/language arts and mathematics components of their assessment systems in place by the 2005 - 2006 school year. These assessments must be administered annually to all students in each of grades 3 - 8 and at least once to students in the 10 - 12 grade range. By the 2007 - 2008 school year, States must also have in place their science assessments, which must be administered, annually, at least once in each of the 3 - 5, 6 - 9, and 10 - 12 grade spans. Assessments administered in the 10 -12 grade range in reading/language arts, mathematics, and science may be end-of-course tests so long as the associated courses, or combinations of courses, are ones that all students must take.

States must also have in place an alternate assessment designed for those students who are unable to participate meaningfully in the regular assessment, even with accommodations. For most students, results from this alternate assessment will be evaluated against the same grade-level achievement standards as the regular test for which it serves as an alternate. With respect to students with the most significant cognitive disabilities, however, the State may define alternate achievement standards. In addition, a few States have developed alternative assessments for use with LEP students. Results from these assessments must be judged against the same grade-level standards as the regular tests.

The assessments that make up the State's assessment system may either be criterion-referenced or an augmented form of a norm-referenced test. If the State uses only assessments referenced against national norms at a particular grade, those assessments must be augmented with additional items as necessary to measure accurately the depth and breadth of the State's academic content standards and express student results in terms of the State's student academic achievement standards.

A State's assessment system may include only statewide assessments, a combination of statewide and local assessments, or only local assessments.¹ However, if the State includes

¹ State law exception as stated in Section 200.4 of the July 5, 2002 Regulations and Section 1111(b)(5) of NCLB.

Standards and Assessments Peer Review Guidance
April 28, 2004

local assessments in its system, the State is responsible for ensuring that each of these assessments meets the rigorous criteria for technical quality and alignment specified in this document. The State must ensure that results from all local assessments can be aggregated meaningfully at the State level with one another and with scores from any statewide assessments.

In building its assessment system, a State must ensure that the information its assessments yield is coherent across grades and content areas. For example, information gained from the reading/language arts assessment at grade 3 should be clearly and appropriately relevant to information gained from the reading/language arts assessment at grade 4 and subsequent grades. This does not require use of tests that are vertically scaled, but does imply the articulation of the standards from grade to grade. The content of the assessments and the achievement standards should be articulated across grades.

**Standards and Assessments Peer Review Guidance
April 28, 2004**

OVERVIEW OF THE STATEWIDE ASSESSMENT SYSTEM

Section 3.1. In the chart below indicate your State’s current assessment system in reading /language arts and mathematics in grades 3 through 8 and for the 10-12 grade range using the abbreviations to show what type of assessments the State’s assessment system is composed of: (a) criterion-referenced assessments (**CRT**); or (b) augmented norm-referenced assessments (**ANRT**) (augmented as necessary to measure accurately the depth and breadth of the State’s academic content standards and yield criterion-referenced scores); or (c) a combination of both across grade levels and/or content areas. Also indicate your current assessment system in science² that is aligned with the State’s challenging academic content and achievement standards at least once in each of the grade spans 3-5, 6-9, and 10-12. A State may have assessments in reading or language arts depending on the alignment to the State’s content standards; both are not required. Please indicate, using the abbreviations shown, the grades and subject areas with availability of native language assessment (**NLA**) or various alternate assessments (**AA-SWD** for an alternate assessment for students with disabilities based on grade-level standards; **AA-LEP** for an alternate assessment for students with limited English proficiency based on grade-level standards, and/or **AA-AAS** for an alternate assessment for students with the most significant cognitive disabilities based on alternate achievement standards).

Chart of State Assessment System Aligned to Content Standards for school year _____ by Subject, Grade, and Type of Assessment

Grades	3	4	5	6	7	8	9	10	11	12
Math										
Alternate										
Native Lang.										
Reading										
Alternate										
Native Lang.										
Language arts										
Alternate										
Native Lang.										
Science										
Alternate										
Native Lang.										

² Science assessments are not due until the 2007-2008 school year.

Standards and Assessments Peer Review Guidance
April 28, 2004

Section 3: Statewide Assessment System

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>3.2 If the State’s assessment system includes assessments developed or adopted at both the local and State level, how has the State ensured that these local assessments meet the same technical requirements as the statewide assessments?</p> <p>(a) How has the State ensured that all local assessments are aligned with the State’s academic content and achievement standards?</p> <p>(b) How has the State ensured that all local assessments are equivalent to one another in terms of content coverage, difficulty, and quality?</p> <p>(c) How has the State ensured that all local assessments yield comparable results for all subgroups?</p> <p>(d) How has the State ensured that all local assessments yield results that can be aggregated with those from other local assessments and with any statewide assessments?</p> <p>(e) How has the State ensured that all local assessments provide unbiased, rational, and consistent determinations of the annual progress of schools and LEAs within the State?</p>	<p>The State has determined that the proficiency-level judgments yielded by each of the local assessments have comparable validity and reliability, are aligned with the State’s academic content standards, are equivalent to one another in content, coverage, difficulty and quality, provide unbiased, rational, and consistent determinations of AYP for schools and LEAs, and may be aggregated at the State level to determine whether the State has made adequate yearly progress.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • The State’s assessment system includes local assessments, written documentation of the State’s criteria for these assessments and the State’s processes for ensuring their quality and comparability. • Documentation of studies conducted by the State or by independent evaluators on the quality and comparability of each of the local assessments as well as follow-up plans for the State to address any deficiencies in these local assessments or their comparability, equivalence, and ability to produce results that can be aggregated. 	<p>The State’s assessment system includes local assessments but the State does not certify that they meet the requirements specified in this document.</p> <p>Results from one or more local assessments do not yield comparable and equivalent results for one or more subgroups.</p> <p>Results from one or more local assessments cannot be meaningfully aggregated with results from the other local assessments for one or more subgroups.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 3: STATEWIDE ASSESSMENT SYSTEM

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>3.3 If the State’s assessment system employs a matrix design—that is, multiple forms within a content area and grade level-- how has the State ensured that:</p> <p>(a) All forms are aligned with the State’s academic content and achievement standards and yield comparable results?</p> <p>(b) All forms are equivalent to one another in terms of content coverage, difficulty, and quality?</p> <p>(c) All assessments yield comparable results for all subgroups?</p>	<p>Possible Evidence</p> <ul style="list-style-type: none"> • Documentation such as technical manuals and studies conducted by the State or by independent evaluators on the quality, equivalence, and comparability of the forms. 	<p>The State does not document the quality, equivalence and comparability of the forms of the assessments.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 3: STATEWIDE ASSESSMENT SYSTEM

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>3.4 How has the State ensured that its assessment system will provide coherent information for students across grades and subjects?</p> <p>(a) Has it indicated the relative contribution of each assessment to ensure alignment to the content standards and determining adequate yearly progress?</p> <p>(b) Has the State provided a rational and coherent design that identifies the assessments to be used?</p>	<p>For each grade and subject assessed, the State’s academic assessment system is designed to provide a coherent system across grades and subjects.</p> <p>The State’s assessments are aligned with the State’s academic achievement standards and reflect articulation of knowledge and skills across grades and subjects.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Existing written documentation describing the processes used to develop the academic content and achievement standards to align the assessments with these standards. • Documentation of the studies used to establish vertical scales that span grade levels within a given content area. • The State identifies the assessments to be used, indicates the relative contribution of each assessment and provides information regarding the progress of students relative to the State’s academic standards in order to inform instruction. • Documentation that indicates how each of the assessments contributes to the alignment to the content standards. • A plan that identifies each of the assessments and their contribution to determining adequate yearly progress. 	<p>The State has not considered or accounted for inconsistencies in the academic achievement standards across grade levels or content areas.</p> <p>The State has not considered or accounted for the appropriate progression of content across grade levels within each content area.</p> <p>The State has not provided a plan that indicates the relative contribution of each of its assessments.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 3: STATEWIDE ASSESSMENT SYSTEM

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>3.5 If its assessment system includes various instruments (e.g., the regular assessment in English and either a native-language version or simplified English version of the assessment), how does the State demonstrate comparable results and alignment with the academic content and achievement standards?</p>	<p>Possible Evidence</p> <ul style="list-style-type: none"> • Documentation of alignment studies of the assessments with the academic content and achievement standards. 	<p>For the State’s various instruments, the State has not demonstrated that these assessments are comparable and are aligned with the content standards.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 3: STATEWIDE ASSESSMENT SYSTEM

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>3.6 How does the State’s assessment system involve multiple measures, that is, measures that assess higher-order thinking skills and understanding of challenging content?</p>	<p>The State’s assessment system ensures coverage of the depth and breadth of its academic content standards and employs multiple approaches within specific grade and content combinations as needed to meet this goal.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Test blueprints or item specifications that describe the structure of each assessment and the items on each form. • Description of the process used to determine and judge the inclusion of challenging content. • Statistical evidence that documents coverage of higher-order thinking skills consistent with the standards. 	<p>The State’s assessment system covers only basic skill items.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 3: STATEWIDE ASSESSMENT SYSTEM

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>3.7 Has the State included alternate assessment(s) for students whose disabilities do not permit them to participate in the regular assessment even with accommodations?</p>	<p>The State’s assessment system includes an alternate assessment that yields a student score for each grade and subject assessed. Results from the alternate assessment may be evaluated against the grade-level standards for the grade in which the student is enrolled or against alternate achievement standards for students with the most significant cognitive disabilities.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Existing written documentation describing the processes used to develop the alternate assessment(s) and the associated achievement standards. • Documentation of the criteria that local IEP teams should use to determine which students are eligible to participate in the alternate assessment. • Sample score reports from the alternate assessment or other materials used to communicate the meaning of results to parents. 	<p>The State has not yet implemented an alternate assessment.</p> <p>The State has adopted an existing standardized test as the alternate assessment with no explanation of how it reflects appropriate standards and procedures for students whose disabilities prevent participation in the regular test.</p> <p>The State has not implemented alternate assessments that meet the requirements under NCLB.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

Section 4: A system of assessments with high technical quality

Reference in NCLB legislation: Sec. 1111(b)(3)
Reference in final regulations: Sec. 200.2

Overview

The *Standards for Educational and Psychological Testing* (1999) delineates the characteristics of high-quality assessments and describes the processes that a State can employ to ensure that its assessments and use of results are appropriate, credible and technically defensible. The *Standards*, developed jointly by the American Psychological Association, the American Educational Research Association, and the National Council of Measurement in Education, has a history of 30 years of use by test developers and the courts.

Validity

As reflected in the *Standards*, the primary consideration in determining validity is whether the State has evidence that the assessment results can be interpreted in a manner consistent with their intended purpose(s).

The *Standards* speaks of four broad categories of evidence used to determine construct validity: (1) evidence based on test content, (2) evidence based on the assessment's relation to other variables, (3) evidence based on student response processes, and (4) evidence from internal structure.

1) Using evidence based on test content (content validity). Content validity, that is, alignment of the standards and the assessment, is important but not sufficient. States must document not only the surface aspects of validity illustrated by a good content match, but also the more substantive aspects of validity that clarify the "real" meaning of a score.

2) Using evidence of the assessment's relationship with other variables. This means documenting the validity of an assessment by confirming its positive relationship with other assessments or evidence that is known or assumed to be valid. For example, if students who do well on the assessment in question also do well on some trusted assessment or rating, such as teachers' judgments, it might be said to be valid. It is also useful to gather evidence about what a test does *not* measure. For example, a test of mathematical reasoning should be more highly correlated with another math test, or perhaps with grades in math, than with a test of scientific reasoning or a reading comprehension test.

3) Using evidence based on student response processes. The best opportunity for detecting and eliminating sources of test invalidity occurs during the test development process. Items obviously need to be reviewed for ambiguity, irrelevant clues, and inaccuracy. More direct evidence bearing on the meaning of the scores can be gathered during the development process by asking students to "think-aloud"

Standards and Assessments Peer Review Guidance

April 28, 2004

and describe the processes they “think” they are using as they struggle with the task. Many States now use this "assessment lab" approach to validating and refining assessment items and tasks.

4) *Using evidence based on internal structure.* A variety of statistical techniques have been developed to study the structure of a test. These are used to study both the validity and the reliability of an assessment. The well-known technique of item analysis used during test development is actually a measure of how well a given item correlates with the other items on the test. Newer technologies including generalizability analyses are variations on the theme of item similarity and homogeneity. A combination of several of these statistical techniques can help to ensure a balanced assessment, avoiding, on the one hand, the assessment of a narrow range of knowledge and skills but one that shows very high reliability, and on the other hand, the assessment of a very wide range of content and skills, triggering a decrease in the consistency of the results.

In validating an assessment, the State must also consider the consequences of its interpretation and use. Messick (1989) points out that these are different functions, and that the impact of an assessment can be traced either to an interpretation or to how it is used. Furthermore, as in all evaluative endeavors, States must attend not only to the intended effects, but also to unintended effects. The disproportional placement of certain categories of students in special education as a result of accountability considerations rather than appropriate diagnosis is an example of an unintended--and negative--consequence of what had been considered proper use of instruments that were considered valid.

Reliability

The term “reliability” is usually defined with synonyms such as consistency, stability, and accuracy. These terms all relate to the problem of uncertainty in making an inference about a score. As reflected in the *Standards for Educational and Psychological Testing*, the field now treats reliability as a study of the many sources of unwanted variation in assessment results. Those responsible for developing and operating State assessment systems are obliged to (1) make a reasonable effort to determine the types of error that may (unwittingly) distort interpretations of the findings, (2) estimate their magnitude, and (3) make every possible effort to alert the users to this lack of certainty.

The traditional methods of portraying the consistency of test results, including reliability coefficients and standard errors of measurement, should be augmented by techniques that more accurately and visibly portray the actual level of accuracy (Rogosa, 1995, Young and Yoon, 1999). Most of these methods focus on error in terms of the probability that a student with a given score, or pattern of scores, is properly classified at a given performance level, such as "proficient." For school-level or district-level results, the report should indicate the estimated amount of error associated with the percent of students classified at each performance level. For example, if a school reported that 47% of its students were proficient, the report might say that the reader could be confident at the 95% level that the school's true percent of students at the proficient level is between 33% and 61%.

Standards and Assessments Peer Review Guidance

April 28, 2004

Furthermore, since the focus on results in a Title I context is on improvement over time, the report should also indicate the accuracy of the year-to-year changes in scores.

Other dimensions of technical quality

There are several other characteristics of State assessments that support valid interpretation and use of results.

Fairness/Accessibility The *Standards* identifies several sources of unfairness, including bias or unequal treatment of students in the assessment process or in the processes of reporting, interpretation, or use; and the lack of opportunity to learn to the standards. Unfairness most often appears at four points in the assessment process:

- The items or tasks do not provide an equal opportunity for all students to fully demonstrate their knowledge and skills.
- The assessments are not administered in ways that ensure fairness.
- The results are not reported in ways that ensure fairness.
- The results are not interpreted or used in ways that leads to equal treatment.

Comparability of results Many uses of State assessment results assume comparability of different types: comparability from year to year, from student to student, and from school to school. Although this is difficult to implement and to document, States have an obligation to show that they have made a reasonable effort to attain comparability, especially where locally selected assessments are part of the system.

Procedures for test administration, scoring, data analysis, and reporting Most States take great pains to ensure that the assessments are properly administered, that directions are followed, and that test security requirements are clearly specified and followed. Nevertheless, it is important they document the ways in which they ensure that their system does not omit any of these basics.

Interpretation and use of results Although this topic is closely related to that of validity, and is discussed in most of the other topics in this section, it is mentioned here because of its importance. Even if an assessment is carefully designed, constructed and implemented, it all can come to naught if users are not helped to draw the most appropriate interpretations and to use the results in the most valid ways.

Validation efforts continue throughout the life of the assessment. Evidence should continually be sought that the results truly reflect the goals of instruction, especially those related to higher-order thinking and understanding. Accurate data about the consequences of an assessment will, obviously, not be available until they have been implemented for a year or more. Research questions might ask: Are more students meeting the standards because the results led to the creation of a dynamic statewide after-school program? Are more students being retained in grade as a result of the assessment results? Are more teachers part of a long-term professional development program that improves the teaching of reading to low-achieving students?

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 4: TECHNICAL QUALITY

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>4.1 For each assessment, including alternate assessment(s), has the State documented the issue of validity (in addition to the alignment of the assessment with the content standards), as described in the <i>Standards for Educational and Psychological Testing</i> (AERA/APA/NCME, 1999), with respect to <u>all</u> of the following categories:</p> <ul style="list-style-type: none"> (a) Has the State specified the purposes of the assessments, delineating the types of uses and decisions most appropriate to each? <u>and</u> (b) Has the State ascertained that the assessments, including alternate assessments, are measuring the knowledge and skills described in its academic content standards and not knowledge, skills, or other characteristics that are not specified in the academic content standards or grade level expectations? <u>and</u> (c) Has the State ascertained that its assessment items are tapping the intended cognitive processes and that the items and tasks are at the appropriate grade level? <u>and</u> (d) Has the State ascertained that the scoring and reporting structures are consistent with the sub-domain structures of its academic content standards (i.e., are item interrelationships consistent with the framework from which the test arises)? <u>and</u> (e) Has the State ascertained that test and item scores are related to outside variables as intended (e.g., scores are correlated strongly with relevant measures of academic achievement and are weakly correlated, if at all, with irrelevant characteristics, such as demographics)? <u>and</u> (f) Has the State ascertained that the decisions based on the results of its assessments are consistent with the purposes for which the assessments were designed? <u>and</u> (g) Has the State ascertained whether the assessment produces intended and unintended consequences? 	<p>For each assessment, including alternate assessment(s), the State has documented the existing validity evidence in each of the categories and has taken steps to address any deficiencies either in validity or in its approach to establishing and documenting validity evidence.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • For category (a), existing written documentation, such as minutes or policies of the State Board of Education or state legislative code, that defines the purpose(s) of the State’s assessment system. • For each of the categories (b) – (g), documentation of the studies that provide evidence in support of the validity of using results from State’s assessment system for their stated purpose(s). 	<p>The State has not provided evidence in all categories (a) -- (g) or has not taken steps to address any deficiencies either in validity or in its approach to establishing and documenting validity evidence.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 4: TECHNICAL QUALITY

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>4.2 For each assessment, including alternate assessment(s), has the State considered the issue of reliability, as described in the <i>Standards for Educational and Psychological Testing</i> (AERA/APA/NCME, 1999), with respect to <i>all</i> of the following categories:</p> <p>(a) Has the State determined the reliability of the scores it reports, based on data for its own student population and each reported subpopulation? <i>and</i></p> <p>(b) Has the State quantified and reported within the technical documentation for its assessments the conditional standard error of measurement and student classification that are consistent at each cut score specified in its academic achievement standards? <i>and</i></p> <p>(c) Has the State reported evidence of generalizability for all relevant sources, such as variability of groups, internal consistency of item responses, variability among schools, consistency from form to form of the test, and inter-rater consistency in scoring?</p>	<p>For each assessment, including alternate assessment(s), the State has documented reliability evidence in each of the categories and has taken steps to address any deficiencies either in reliability or in the State’s approach to establishing and documenting reliability evidence.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • For each of the categories (a) – (c), documentation of the studies that support the reliability of each of the State’s assessments with the State’s own student population. • Documentation of the precision of the assessments at cut scores and evidence of a systematic process for addressing any deficiencies identified in these studies. • Documentation of consistency of student level classification and evidence of a systematic process for addressing any deficiencies identified in these studies. 	<p>The State has not provided evidence in all categories (a) -- (c) or has not taken steps to address any deficiencies either in reliability or in the State’s approach to establishing and documenting reliability evidence.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 4: TECHNICAL QUALITY

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>4.3 Has the State ensured that its assessment system is fair and accessible to all students, including students with disabilities and students with limited English proficiency, with respect to each of the following issues:</p> <p>(a) Has the State ensured that the assessments provide an appropriate variety of accommodations for students with disabilities? <i>and</i></p> <p>(b) Has the State ensured that the assessments provide an appropriate variety of linguistic accommodations for students with limited English proficiency? <i>and</i></p> <p>(c) Has the State taken steps to ensure fairness in the development of the assessments? <i>and</i></p> <p>(d) Does the use of accommodations and/or alternate assessments yield meaningful scores?</p>	<p>The State has taken appropriate judgmental (e.g., committee review) and data-based (e.g., bias studies) steps to ensure that its assessment system is fair and accessible to all students. Review committees have included representation of identified subgroups.</p> <p>The State assessment system must be designed to be valid and accessible for use by the widest possible range of students.</p> <p>The State is conducting studies to determine the appropriateness of accommodations and the impact on test scores.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Existing written documents describe how the principles of universal design and/or appropriate language simplification were incorporated into each of the State’s assessments. • Evidence that students with disabilities were included in the test development process. • Existing written documentation of the State’s policies and procedures for the selection and use of accommodations and alternate assessments, including evidence of training for educators who administer these assessments. 	<p>The State has conducted data-based bias studies but has not convened committees of stakeholders to review its assessment items.</p> <p>The State has convened committees of stakeholders to review its assessment items but these committees have not included representation of identified subgroups.</p> <p>The State assessment system is not designed to be valid and accessible for use by the widest possible range of students.</p> <p>The State does not have a policy on the appropriate selection and use of accommodations and alternate assessments.</p> <p>The State does not train or monitor personnel at the school, LEA, and State levels with regard to the appropriate selection and use of accommodations and alternate assessments.</p> <p>There are no appropriate accommodations for students with particular disabilities (e.g., no allowable accommodations on the regular assessment or alternate assessments for students who are visually impaired and need large print or Braille or for students who are significantly physically impaired and need assistive technology.)</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 4: TECHNICAL QUALITY

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>4.4 When different test forms or formats are used, the State must ensure that the meaning and interpretation of results are consistent.</p> <p>(a) Has the State taken steps to ensure consistency of test forms over time?</p> <p>(b) In the State administers both an online and paper and pencil test, has the State documented the comparability of the electronic and paper forms of the test?</p>	<p>The State has conducted appropriate equating or linking studies and has presented data that support the success of the equating or linking.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Documentation describing the State’s approach to ensuring comparability of assessments and assessment results across groups and time. • Documentation of equating studies that confirm the comparability of the State’s assessments and assessment results across groups and across time, as well as follow-up documentation describing how the State has addressed any deficiencies. 	<p>The State has not conducted or documented equating studies to establish whether test forms are comparable across time.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 4: TECHNICAL QUALITY

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>4.5 Has the State established clear criteria for the administration, scoring, analysis, and reporting components of its assessment system, including alternate assessment(s) and does the State have a system for monitoring and improving the on-going quality of its assessment system?</p>	<p>The State developed a set of management controls or standards for each of these components and has communicated these criteria to its contractor(s), LEAs, and schools. It requires its contractor(s) to provide specific information on the degree to which each criterion is met.</p> <p>The State uses an extensive system of training and monitoring to ensure that each person who is responsible for handling or administering any portion of its assessments does so in a way that protects the security of the assessments and maintains equivalence of administration conditions across students and schools.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • The State’s criteria for administration, scoring, analysis, and reporting are communicated to its contractor(s). • The State’s test security policy and consequences for violation are communicated to the public and to local educators. • Existing written documentation of the State’s plan for training and monitoring assessment administration conditions across the State, even when its assessment system is comprised of only local assessments. • Documentation that the tests clearly delineate which accommodations may be used for specific sections of the test (e.g., specify the items/sections for which a calculator may be used without invalidating the test). 	<p>The State does not have a test security policy.</p> <p>The State does not train or monitor personnel at the school, LEA, and State levels with regard to its test administration procedures and security policy.</p> <p>The State provides no criteria to its contractor(s) regarding the quality control and security measures it requires for its assessment system.</p> <p>The State provides no criteria to its contractor(s) to ensure that the procedures for scoring of open-ended tasks meet industry standards for accuracy.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 4: TECHNICAL QUALITY

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>4.6 Has the State evaluated its use of accommodations?</p> <p>(a) How has the State ensured that appropriate accommodations are available to students with disabilities and that these accommodations are used in a manner that is consistent with instructional approaches for each student, as determined by a student’s IEP or 504 plan?</p> <p>(b) How has the State determined that scores for students with disabilities that are based on accommodated administration conditions will allow for valid inferences about these students’ knowledge and skills and can be combined meaningfully with scores from non-accommodated administration conditions?</p> <p>(c) How has the State ensured that appropriate accommodations are available to limited English proficient students and that these accommodations are used as necessary to yield accurate and reliable information about what limited English proficient students know and can do?</p> <p>(d) How has the State determined that scores for limited English proficiency students that are based on accommodated administration circumstances will allow for valid inferences about these students’ knowledge and skills and can be combined meaningfully with scores from non-accommodated administration circumstances?</p>	<p>The State provides for the use of appropriate accommodations and has conducted studies to ensure that scores based on accommodated administrations can be meaningfully combined with scores based on the standard administrations.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • The State has analyzed the use of specific accommodations for different groups of students with disabilities and has provided training to support sound decisions by IEP teams. • The State routinely monitors the extent to which test accommodations are consistent with those provided during instruction. • The State has analyzed the effect of specific accommodations for students with limited English proficiency and has shared results with LEAs and schools. • Documentation of the quality and consistency of the accommodations it offers for limited English proficient students (e.g., training of translators, simplified English, standardized translation of instructions for test administration that are comparable to the regular assessment). 	<p>No analyses have been carried out to determine whether specific accommodations produce the effect intended.</p> <p>The State does not require that decisions about how students with disabilities will participate in the assessment system be made on an individual basis or specify that these decisions must be consistent with the routine instructional approaches as identified by each student’s IEP and/or 504 plan.</p> <p>The State uses the same accommodations for limited English proficient students as it uses for students with disabilities.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

Section 5: Alignment of Academic Content Standards, Academic Achievement Standards, and Assessments

Reference in NCLB legislation:	Sec. 1111(b)(1) and 1111(b)(3)
Reference in final regulations:	Sec. 200.2 and 200.3

Overview

A State's system of standards and assessments will provide useful information for valid accountability decisions and educational improvement only to the extent that all components of this system are aligned. If a State's assessments do not adequately measure the knowledge and skills specified in the State's academic content standards, or if they measure something other than what these standards specify, it will be difficult to determine whether students have achieved the intended knowledge and skills. As a result, it will be difficult to make appropriate policy, program, and instructional decisions meant to improve students' achievement. Further, if a State's assessments do not include items that cover the full range of the State's academic achievement standards, it may be difficult to determine whether students have reached the level of proficiency these standards describe.

Alignment encompasses several dimensions; demonstrating that an assessment system is aligned with a State's standards requires more than simply determining whether all the items on the assessment can be matched to one or more standards or whether each of the academic content standards can be matched to one or more items in the assessments. *Alignment is more than this two-way process.* To ensure that its standards and assessments are aligned, a State needs to consider whether the assessments--

- Cover the full range of content specified in the State's academic content standards, meaning that all of the standards are represented legitimately in the assessments; *and*
- Measure both the content (what students know) and the process (what students can do) aspects of the academic content standards; *and*
- Reflect the same degree and pattern of emphasis apparent in the academic content standards (e.g., if the academic content standards place a lot of emphasis on operations then so should the assessments); *and*
- Reflect the full range of cognitive complexity and level of difficulty of the concepts and processes described, and depth represented, in the State's academic content standards, meaning that the assessments are as demanding as the standards; *and*
- Yield results that represent all achievement levels specified in the State's academic achievement standards.

In addition to considering each of these aspects of alignment through a systematic development and review process, the State needs to also develop strategies for communicating to its education stakeholders how its standards and assessment are aligned. Parents, educators, and other stakeholders need to know how assessment results are related to content-based expectations in order to understand and use test information effectively.

Standards and Assessments Peer Review Guidance
April 28, 2004

Each State must present evidence that its assessment system is aligned to its standards. Some alignment evidence is generated in the test development process, and documentation of the steps taken to ensure that items were drafted to reflect the full range of the State standards is appropriate verification of efforts to attain alignment. In addition, final alignment of assessments and standards following full implementation should be confirmed using one of several procedures (for example, review and comment by external subject-matter experts). Occasionally, documentation of alignment includes the process of re-verification if changes in tests were made to improve alignment.

In recent years, several methods of evaluating alignment between standards and assessments have been developed. A summary and comparison of alignment models can be found on the Council of Chief State Officers website at:

http://www.ccsso.org/Projects/alignment_analysis/models/418.cfm

When documenting the comprehensive aspects of alignment between standards and the State assessment system, the State should describe--

- The relationships between the structure of the standards and the structure of the assessments;
- The rationale for the overall alignment strategy, including a rationale for any standards either not assessed or not reported as part of the State assessment;
- The manner in which each standard is assessed, whether at the State, district, school, or classroom level;
- The manner in which alternate assessments based on alternate achievement standards are linked to the State content standards; and
- The type of information the State collects pertaining to each standard, and how the State monitors the quality of the assessment data collected at the local level, for all assessments that are part of the statewide system.

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 5: ALIGNMENT

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>5.1 Has the State outlined a coherent approach to ensuring alignment between each of its assessments, including alternate assessment(s), or combination of assessments, and the academic content standards and academic achievement standards the assessment is designed to measure?</p>	<p>The State has developed an assessment system consistent with its academic content and achievement standards and is implementing on-going quality control reviews to ensure that the system remains fully aligned over time.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Detailed assessment specifications and a description of the process used to ensure that full alignment is achieved initially and maintained over time through quality control reviews. • Descriptions of the internal and external groups involved in the State’s alignment process. • Reports of independent alignment studies (i.e., conducted by an entity other than the State or its assessment contractor) and evidence of a systematic process for addressing any gaps or weaknesses identified in these studies. • If the State has multiple assessments within one grade level in reading/language arts or mathematics, then the State has tapped all content sub-domains. The State is implementing a series of studies to ensure that this combination is aligned to the full scope of the domain. 	<p>The State accepts its contractor’s assurance as its sole evidence of alignment.</p> <p>The State has studied whether all of the items on its assessments match its academic content standards but has not conducted studies to ensure that all of its academic content standards are reflected by items on its assessments.</p> <p>The State has conducted alignment studies for some, but not all, of the assessments in its system.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 5: ALIGNMENT

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>5.2 Are the assessments and the standards aligned comprehensively, meaning that the assessments reflect the full range of the State’s academic content standards? Are the assessments as cognitively challenging as the standards? Are the assessments and standards aligned to measure the depth of the standards? Do the standards reflect the degree of cognitive complexity and level of difficulty of the concepts and processes described in the standards?</p>	<p>The State’s assessment plan, assessment blueprints, and/or item/task specifications describe how all content standards are assessed and how the domain is sampled to lead to valid inferences about student performance on the standards, individually and in the aggregate. The State has evidence that (a) the full scope of the standards and their differential emphases are reflected in the plan/blueprints/specifications and that (b) the assessments match the plan/blueprints/specifications. Impartial experts were involved in this process.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Detailed assessment specifications and a description of the process that was used to ensure that its assessment system reflects the full range of content and level of challenge specified in its academic content standards, as well as the range of performance indicated in its academic achievement standards. • Reports of independent alignment studies and evidence of a systematic process for addressing any gaps or weaknesses identified in these studies. 	<p>The State makes an assertion of comprehensiveness without documentation matching both assessments to standards and standards to assessments.</p> <p>The State’s assessments do not appear to measure the more challenging aspects of its standards.</p> <p>The State’s assessment items measure higher-order thinking, but do not measure all of the standards that call for higher-order thinking.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 5: ALIGNMENT

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>5.3 Are the assessments and the standards aligned in terms of both content (knowledge) and process (how to do it), as necessary, meaning that the assessments measure what the standards state students should both know and be able to do?</p>	<p>The State’s assessments reflect both the content and the process dimensions of the academic content standards. These assessments are designed in a way that will allow students to demonstrate content knowledge through activities described in the standards.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Detailed assessment specifications and a description of the process used to ensure that its assessment system reflects both the content and the processes and skills specified in its academic content standards. • Reports of independent alignment studies and evidence of a systematic process for addressing any gaps or weaknesses identified in these studies. 	<p>Items on the State’s assessments address only content dimensions of the State’s standards and not the process or skill dimensions in these standards.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 5: ALIGNMENT

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>5.4 Do the assessments reflect the same degree and pattern of emphasis as are reflected in the State’s academic content standards?</p>	<p>The number of score points in content sub-domains on the State’s assessment is consistent with the representation of these sub-domains in the State’s academic content standards.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Detailed assessment specifications and a description of the process used to ensure that its assessment system reflects the degree and patterns of emphasis that are specified in its academic content standards. • Reports of independent alignment studies and evidence of a systematic process for addressing any gaps or weaknesses identified in these studies. 	<p>One or more sub-domains in the State’s academic content standards are under- or over-represented by score points on its assessments.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 5: ALIGNMENT

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>5.5 Do the assessments yield scores that reflect the full range of achievement implied by the State’s academic achievement standards?</p>	<p>The State’s assessments have sufficient items at each level to permit students to demonstrate the full range of the State’s academic achievement standards.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Detailed assessment specifications and a description of the process used to ensure that its assessment system reflects the full range of achievement described in its academic achievement standards • Reports of independent alignment studies and evidence of a systematic process for addressing any gaps or weaknesses identified in these studies. 	<p>The items do not reflect the full range of achievement implied by the State’s academic achievement standards.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 5: ALIGNMENT

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>5.6 Assessment results must be expressed in terms of the achievement standards, not just scale scores or percentiles.</p>	<p>The State has designed reports and communicated assessment results in terms of its achievement standards.</p> <p>The State’s assessments yield scores that are clearly aligned with the State’s academic content standards at the domain and/or sub-domain levels.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Examples of existing documents, such as web pages, brochures, guidelines, or media reports, designed to communicate the alignment between the standards and assessments to all members of the school community. • “Extended” standards communicate the relationship between the State’s academic content standards and the content of the alternate assessment based on alternate achievement standards. 	<p>The State provides no information about the alignment of its standards and assessments for educators, parents, or the public. Results are expressed only as percentiles or normal curve equivalents.</p> <p>The State indicates or implies that there really is no easy way for teachers or the public to see whether or how well the assessments are aligned with the standards.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

SECTION 5: ALIGNMENT

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>5.7 What ongoing procedures does the State use to maintain and improve alignment between the assessment(s) and standards over time?</p>	<p>The State has used the information gained through its series of alignment studies to eliminate gaps and weaknesses in alignment and is implementing a plan for continuous quality review to maintain alignment over time.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Documentation of independent alignment studies. • If any independent alignment studies reveal gaps or weaknesses in the alignment of the State’s assessments and standards, existing written documentation describing the State’s systematic process for addressing these deficiencies. 	<p>The State has not implemented strategies, such as adding items to the assessment, adding multiple measures, adding a writing test, or adopting the longer version of a test, to address the gaps and weaknesses identified in its alignment studies.</p>

List State Evidence Here

Standards and Assessments Peer Review Guidance
April 28, 2004

Section 6: Inclusion of all students in the assessment system

Reference in NCLB legislation: Sec. 1111(b)(3)
Reference in final regulations: Sec. 200.6

Overview

Just as its title indicates, one of the fundamental principles of the NCLB is the inclusion of *all* students in a state's system of standards, assessments, and accountability. By excluding any student or group of students from its assessment system, a state suggests that its high expectations apply only to some, but not all, students.

For some students with disabilities and for students who are not yet proficient in English, participation in the State's assessment system may require special considerations.³ In all cases, however, decisions must be made regarding how an individual student will participate in the assessment system, not whether the student will participate.

To ensure that all students can participate fully in its assessment system, a State must allow:

- Participation in the regular assessment (limited English proficient students and students with disabilities); and
 - Participation in the regular assessment through the use of one or more approved accommodations (limited English proficient students and students with disabilities);
 - At least one alternate assessment, which may involve either or both of the following:
 - Participation in an alternate assessment that is aligned with the State's academic content standards and based on grade-level achievement standards (limited English proficient students and students with disabilities);
- and/or
- Participation in an alternate assessment that is based on alternate achievement standards (limited to students with the most significant cognitive disabilities).

Implementation of these options will require States to identify the needs of its special student populations so that it can appropriately address these needs. For example, for students who are visually- or hearing-impaired, the State needs to make available appropriate accommodations that will allow these students to demonstrate what they know and can do, as well as develop a system for ensuring that these accommodations are selected and used appropriately. For students with disabilities who cannot participate in the State's regular assessments, even with accommodations, the State must offer an alternate assessment that is based on the State's academic content standards, yields results in both

³ Letter from Secretary Rod Paige dated February 20, 2004 on the flexibility in assessing new limited English proficient students and in measuring adequate yearly progress.

Standards and Assessments Peer Review Guidance
April 28, 2004

reading/language arts and mathematics, and is designed and implemented in a manner that supports use of the results as an indicator of adequate yearly progress. For students with limited English proficiency, the State must offer accommodations including, to the extent practicable, assessments designed to ensure that these students have an opportunity to demonstrate their academic knowledge and skills based on grade-level standards.

In addition to addressing the needs of students with disabilities and students with limited English proficiency, a State must take steps to ensure the participation of all migrant, otherwise mobile, and homeless students in its assessment system. This includes the accurate identification of migrant students and policies requiring assessment of all students, regardless of how long these students have been enrolled in the State.

It is important to note that as States continue to improve alignment between standards and assessments, the use of universal design principles holds great promise for designing and aligning standards, curriculum, instructional materials and strategies. Assessments that are designed to be valid and accessible for the widest possible range of students may help all students struggling to achieve, particularly students with cognitive disabilities, and would reduce the need for accommodations.

Standards and Assessment Peer Review Guidance
April 28, 2004

SECTION 6: INCLUSION

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>6.1 Do the State’s participation data indicate that all students in the tested grade levels or grade ranges are included in the assessment system (e.g., students with disabilities, students with limited English proficiency, economically disadvantaged students, race/ethnicity, migrant students, homeless students, etc.)?</p>	<p>The State has documented its total and subgroup enrollments in each of the required grade levels or grade ranges and calculates its participation rates as a proportion of students assessed to students enrolled during the test administration period.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Reports that specify the participation rates and the method of calculations for all students and for each subgroup in the assessment system. 	<p>The State does not offer participation data for all students.</p> <p>The State provides participation data for assessments that do not meet NCLB requirements.</p>

List State Evidence Here

Standards and Assessment Peer Review Guidance
April 28, 2004

SECTION 6: INCLUSION

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>6.2</p> <p>1. What guidelines does the State have in place for including all students with disabilities in the regular assessment system?</p> <p>2. If the State has approved/adopted alternate achievement standards, what guidelines does the State have in place for assessing only students with the most significant cognitive disabilities based on alternate achievement standards?</p> <p>(a) Has the State developed clear guidelines for Individualized Educational Program (IEP) teams to apply in determining when a child’s cognitive disability justifies assessment based on alternate academic achievement standards?</p> <p>(b) Has the State ensured that parents of those students are informed that their child’s achievement will be based on alternate achievement standards?</p> <p>(c) Has the State documented that students with the most significant cognitive disabilities are, to the extent possible, included in the general curriculum and assessments aligned to that curriculum?</p> <p>(d) Has the State developed, disseminated information on, and promoted use of appropriate accommodations to increase the number of students with the most significant cognitive disabilities who are tested against grade-level academic achievement standards?</p> <p>(e) Has the State ensured that regular and special education teachers and other appropriate staff know how to administer assessments, including making use of accommodations, for students with the most significant cognitive disabilities?</p>	<p>The State provides clear, written guidelines to all LEAs concerning how to appropriately include <i>all</i> students with disabilities in the assessment system. This statement specifies that decisions about how to include students with disabilities (i.e., whether a student takes the regular assessment without accommodations, the regular assessment with accommodations, an alternate assessment based on grade-level achievement standards, or an alternate assessment based on alternate achievement standards) must be made on an individual basis. Further, decisions about how a student will participate must be consistent with the routine instructional approaches as identified by each student’s IEP or 504 plan.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • The State’s guidelines, as communicated to LEAs, for the inclusion of all students with disabilities in the assessment system. • Existing written documentation describing the form and content of alternate assessments for students with disabilities, the process by which these assessments were developed, and the process by which the State has ascertained the alignment of these assessments with its academic content standards and academic achievement standards—both grade-level and alternate, as applicable. • Report that shows that 100% of students with 	<p>The State does not provide clear, written guidelines to all LEAs on how to maximize inclusion of students with disabilities in its regular assessment system.</p> <p>The State allows some students with disabilities to be exempted from participating in the assessment system.</p> <p>The State assessment system does not include an alternate assessment.</p>

Standards and Assessment Peer Review Guidance
April 28, 2004

	<p>disabilities are taking the regular assessment or an alternate assessment.</p> <ul style="list-style-type: none">• Curriculum guides that inform educators about the inclusion of students with disabilities in regular assessments, with or without accommodations, or alternate assessments based on grade-level standards, or, if the State allows it, alternate assessments based on alternate achievement standards.	
--	--	--

List State Evidence Here

Standards and Assessment Peer Review Guidance
April 28, 2004

SECTION 6: INCLUSION

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>6.3 What guidelines does the State have in place for including all students with limited English proficiency in the tested grades in the assessment system?⁴</p> <p>(a) Has the State made available assessments, to the extent practicable, in the language and form most likely to yield accurate and reliable information on what these students know and can do?</p> <p>(b) Does the State require the participation of every limited English proficient student in the assessment system, regardless of how long a student has been enrolled in US schools? Has the State adopted policies requiring limited English proficient students to be assessed on the reading/language arts standards in English if they have been enrolled in US schools for three consecutive years or more?</p>	<p>The State provides clear, written guidelines to all LEAs concerning how to include <i>all</i> limited English proficient students in the assessment system. These guidelines specify that decisions about how to include limited English proficient students (i.e., whether a student takes the regular assessment without accommodations, or the regular assessment with accommodations) must be made on an individual basis.</p> <p>The State provides an alternate assessment in English aligned with grade-level standards for limited English proficient students who have not yet acquired a level of proficiency in English that would allow them to participate in the regular assessments, even with accommodations.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • The State’s guidelines, as communicated to LEAs, calls for the inclusion of all limited English proficient students in its assessment system. • Existing written documentation describing the form and content of any alternate assessments for limited English proficient students, the process by which these assessments were developed, and the process by which the State has ascertained the alignment of these assessments with its academic content standards 	<p>The State does not provide clear, written guidelines to all LEAs on how <i>all</i> limited English proficient students are to be included in its assessment system.</p> <p>The State allows some limited English proficient students to be exempted from participating in the assessment system.</p> <p>The State allows limited English proficient students who have not been enrolled in its school system for at least 1 year to be categorically exempted from participation in the assessment system.</p> <p>The State does not require decisions about how limited English proficient students will participate in the assessment system to be made on an individual basis.</p> <p>The State does not offer assessments for limited English proficient students who have not yet acquired a level of proficiency in English that would allow them to participate in the regular assessments, even with accommodation.</p>

⁴ Letter from Secretary Rod Paige dated February 20, 2004 on the flexibility in assessing new limited English proficient students and in measuring adequate yearly progress. For assessments administered during the 2003-04 school year, including those given during the fall 2003, the Secretary has permitted States to exempt recently arrived LEP students (i.e., those students in their first year of enrollment in U. S. schools) from taking the State’s reading/language arts assessment.

Standards and Assessment Peer Review Guidance
April 28, 2004

	and academic achievement standards and comparability of results with the regular test.	
--	--	--

List State Evidence Here

Standards and Assessment Peer Review Guidance
April 28, 2004

SECTION 6: INCLUSION

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>6.4 What policies and practices does the State have in place to ensure the identification and inclusion of migrant and other mobile students in the tested grades in the assessment system?</p>	<p>The State provides clear, written guidelines to all LEAs concerning how to identify and include <i>all</i> migrant and other mobile students in its assessment system.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • The State’s guidelines, as communicated to LEAs, for the inclusion of all migrant and other mobile students in its assessment system. 	<p>The State allows some migrant or other mobile students to be exempted from participating in its assessment system.</p> <p>The State does not have a valid and reliable method for identifying migrant students.</p>

List State Evidence Here

Standards and Assessment Peer Review Guidance
April 28, 2004

Section 7: An effective system of assessment reports

Reference in NCLB legislation: Sec. 1111(b)(3)
Reference in final regulations: Sec. 200.8

Overview

A State's assessment reports represent the culmination of all other aspects of its standards and assessment system. In these reports, a parent, educator, or other stakeholder should find answers to questions about how well a student or group of students is achieving, as well as important information on how to improve achievement in the future.

NCLB requires States to produce reports at the individual student, school, LEA, and State levels. At each of these levels, reports must include scores that are aligned with the State's academic content standards. Also, total test scores must be reported in relation to the performance levels defined in the State's academic achievement standards

Each of a State's reports should be produced and disseminated as soon as possible after each assessment administration. The individual student reports, at least, also need to be accompanied by interpretive guidance that will help parents and educators understand and be able to use the information the reports provide. States must ensure that this guidance is accessible to all parents.

States must carefully protect the data files containing student-level information that are produced following each assessment administration. When the State allows access to this information, it must do so in a way that maintains the confidentiality of each student's records.

Standards and Assessment Peer Review Guidance
April 28, 2004

SECTION 7: ASSESSMENT REPORTS

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>7.1 Does the State’s reporting system facilitate appropriate, credible, and defensible interpretation and use of its assessment data?</p>	<p>The State’s reporting system includes supporting information to facilitate accurate interpretation of data for those who will receive and use its reports, such as information about the content and structure of its assessments, and how the assessments are related to its standards. The State uses a variety of ways to publicize this information, such as manuals, bulletins, reports of results, and websites.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Examples of the State’s score reports at the individual student, school, LEA, and State levels. • Examples of the interpretive guides that accompany reports. • Descriptions of the State’s system for training educators on the appropriate interpretation and use of assessment results. State training materials include: the purpose and content of the assessments, the reliability of the assessment scores, and sufficient information to allow use of the assessment results in making sound educational decisions or for conducting scientifically based research to improve educational outcomes. 	<p>The State’s reporting system does not include training for those who will receive and use its reports on the appropriate interpretation and use of its assessment results.</p> <p>The State does not include interpretive guidance within or attached to each of its assessment reports or this guidance is not easy for stakeholders to access, understand and use.</p>

List State Evidence Here

Standards and Assessment Peer Review Guidance
April 28, 2004

SECTION 7: ASSESSMENT REPORTS

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>7.2 Does the State report participation and assessment results for all students and for each of the required subgroups in its reports at the school, LEA, and State levels? In these assessment reports, how has the State ensured that assessment results are not reported for any group or subgroup when these results would reveal personally identifiable information about an individual student?</p>	<p>The State reports participation and performance results for all students and for each required subgroup at the school, LEA, and State levels. The State has established and justified the minimum number and minimum and maximum proportions of students necessary to allow reporting of scores for any group or subgroup to ensure that personally identifiable information about any individual student is not reported publicly.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Examples of assessment score reports at the school, LEA, and State levels. • Documentation describing the State’s rules for determining whether data are reported for a group or subgroup as well as a description of how these rules are implemented and monitored. • Documentation that electronic student files from the SEA are available to LEAs. 	<p>The State does not disaggregate and report scores for one or more required subgroups even when these subgroups are relatively large.</p> <p>The State reports all scores, regardless of the size of the subgroup.</p>

Standards and Assessment Peer Review Guidance
April 28, 2004

SECTION 7: ASSESSMENT REPORTS

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>7.3 How has the State provided for the production of individual interpretive, descriptive, and diagnostic reports following each administration of its assessments?</p> <p>(a) Do these individual student reports provide valid and reliable information regarding achievement on the assessments in relation to the State’s academic content and achievement standards?</p> <p>(b) Do these individual student reports provide information for parents, teachers, and principals to help them understand and address a student’s specific academic needs? Is this information displayed in a format and language that is understandable to parents, teachers, and principals and are the reports accompanied by interpretive guidance for these audiences?</p> <p>(c) How has the State ensured that these individual student reports will be delivered to parents, teachers, and principals as soon as possible after the assessment is administered?</p>	<p>For each student who participates in the assessment system, the State disseminates two or more copies of an individual student report to the student’s school as soon as possible after each assessment administration. One of these copies is sent to the student’s home by the school and at least one copy is kept in the student’s files. The scores in this report reflect performance in domains and subdomains defined in the State’s academic content standards and indicate which of the achievement levels the student’s scores correspond to. The State includes interpretive guidance with each of the individual student reports and supports local efforts to translate this guidance as needed to make it accessible to parents who do not read English. The guidance includes information about the reliability of the scores that are reported.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Examples of the State’s individual student reports for each grade and content area combination. • Examples of the interpretive guidance that is designed to accompany these reports, including information about how this guidance is made accessible to all parents. • Documentation of the scoring and reporting timeline for each assessment. 	<p>The State does not provide individual student reports for each participating student following each administration of its assessments.</p> <p>Scores on the State’s individual student reports are reported only at the total test level or otherwise are not aligned with the domains and subdomains defined in the State’s academic content standards.</p> <p>Scores on the State’s individual student reports reflect only overall means or percentile ranks or are otherwise not directly associated with the State’s academic achievement standards.</p> <p>The State does not include information about the reliability of the scores reported in the individual student reports in the guidance that accompanies these reports.</p> <p>The State does not provide adequate interpretive guidance to accompany its individual score reports or this guidance is overly complex or not accessible to parents who do not read English.</p> <p>The State’s individual student reports are not delivered to parents, teachers, or principals as soon as possible following each administration of its assessments.</p> <p>Student results are available only through electronic media and therefore not readily available to all parents.</p>

List State Evidence Here

Standards and Assessment Peer Review Guidance
April 28, 2004

SECTION 7: ASSESSMENT REPORTS

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>7.4 How has the State ensured that student-level assessment data are maintained securely to protect student confidentiality?</p>	<p>The State has a clear policy and detailed procedures for allowing access to its student-level assessment data. The State stores these data in a manner that is secure both physically and electronically.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Documentation of the State’s policies and procedures for allowing access to its student-level data files. 	<p>The State posts student-level data on an unsecured website.</p> <p>The State allows liberal access to its student-level assessment data or retains students’ names or other variables that could be used to identify a particular student in the files that it allows to be used for research or evaluation purposes.</p>

List State Evidence Here

Standards and Assessment Peer Review Guidance
April 28, 2004

Section 7: Assessment Reports

Critical Element	Examples of Acceptable Evidence	Examples of Incomplete Evidence
<p>7.5 How has the State provided for the production of itemized score analyses so that parents, teachers, and principals can interpret and address the specific academic needs of students?</p>	<p>In its reports at the student, classroom, school, and LEA levels, the State includes results for each of its academic content standards and also each of the subdomains/ strands within these standards, to the extent that these subscores are based on enough items or score points to be meaningful.</p> <p>Possible Evidence</p> <ul style="list-style-type: none"> • Examples of assessment score reports at the individual student, school, LEA, and State levels; examples of classroom-level reports if the State produces them. 	<p>The State provides only total test scores.</p> <p>The State provides subdomain scores on LEA reports but not on school reports.</p>

List State Evidence Here