U. S. Department of Education Peer Review of State Assessment Systems

Non-Regulatory Guidance for States

for Meeting Requirements of the Elementary and Secondary Education Act of 1965, as amended

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

September 25, 2015

The Department has determined that this document is a "significant guidance document" under the Office of Management and Budget's Final Bulletin for Agency Good Guidance Practices, 72 Fed. Reg. 3432 (Jan. 25, 2007), available at www.whitehouse.gov/sites/default/files/omb/fedreg/2007/012507_good_guidance.pdf. The purpose of this guidance is to provide States, with information to assist them in meeting their obligations under Title I of the Elementary and Secondary Education Act of 1965, as amended. This guidance does not impose any requirements beyond those required under applicable law and regulations. It does not create or confer any rights for or on any person. If you are interested in commenting on this guidance, please e-mail Monique Chism at Monique.Chism@ed.gov or write to us at the following address: U.S. Department of Education, Office of Elementary and Secondary Education, 400 Maryland Avenue, SW, Washington, D.C. 20202.

Paperwork Burden Statement

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless such collection displays a valid OMB control number. The valid OMB control number for this information collection is 1810-0576.

TABLE OF CONTENTS

I – ASSESSMENT PEER REVIEW PROCESS

A. INTRODUCTION

Purpose

The purpose of the U.S. Department of Education's (Department) peer review of State assessment systems is to support States in meeting statutory and regulatory requirements under Title I of the Elementary and Secondary Education Act of 1965 (ESEA) for implementing valid and reliable assessment systems and, where applicable, provide States approved for ESEA flexibility with an opportunity to demonstrate that they have met requirements for high-quality assessments under Principle 1 of ESEA flexibility. Under section 1111(e) of the ESEA and 34 C.F.R. § 200.2(b)(5), the Department also has an obligation to conduct peer reviews of State assessment systems implemented under section 1111(b)(3) of the ESEA.[1]

This guidance is intended to support States in developing and administering assessment systems that provide valid and reliable information on how well students are achieving a State's challenging academic standards to prepare all students for success in college and careers in the 21st century. Additionally, it is intended to help States prepare for assessment peer review of their assessment systems and help guide assessment peer reviewers who will evaluate the evidence submitted by States. The guidance includes: (1) information about the assessment peer review process; (2) instructions for preparing evidence for submission; and (3) examples of evidence for addressing each critical element.

Background

A key purpose of Title I of the ESEA is to promote educational excellence and equity so that by the time they graduate high school all students master the knowledge and skills that they need in order to be successful in college and the workforce. States accomplish this, in part, by adopting challenging academic content standards that define what the State expects all students to know and be able to do. States must develop and administer assessments aligned to those standards, and adopt academic achievement standards aligned to the academic content standards to define levels of student achievement on the assessments.

Specifically, under Title I, each State is responsible for implementing a State assessment system that is coherent and consistent within the State. Section 1111(b)(1) of the ESEA requires a State to develop and implement challenging academic content and achievement standards in at least mathematics, reading/language arts, and science, and to apply the same academic standards to all public schools and public school students in the State. Under section 1111(b)(3) of the ESEA, a State must annually administer State-determined assessments in reading/language arts and mathematics in each of grades 3 through 8 and once in high school, and must annually administer

[1] ESEA Title I, Part A—Improving Basic Programs Operated by Local Educational Agencies, Subpart 1—Basic Program Requirements, 'Section 1111—State Plans is available at: www.ed.gov/policy/elsec/leg/esea02/pg2.html; corresponding regulations, 34 C.F.R. 200.1 – 200.10 —Title I—Improving the Academic Achievement of the Disadvantaged, Subpart A—Improving Basic Programs Operated by Local Educational Agencies, Standards and Assessments, are available at: www.ecfr.gov/cgi-bin/text-idx?SID=084510bc70b2783da282d5f81e8db23e&node=34:1.2.2.1.1.1.145&rgn=div7.

State-determined assessments in science at least once in each of three grade spans (3-5, 6-9 and 10-12). The assessments must be aligned with the full range of the State's academic content standards; be valid, reliable, and of adequate technical quality for the purposes for which they are used; express student results in terms of the State's student academic achievement standards; and provide coherent information about student achievement. The same assessments must be used to measure the achievement of all students in the State, including English learners and students with disabilities, with the exception allowed under 34 C.F.R. Part 200 of students with the most significant cognitive disabilities who may take an alternate assessment based on alternate academic achievement standards.

Within the parameters noted above, each State has the responsibility to design its assessment system. This responsibility includes the adoption of specific academic content standards and State selection of specific assessments. Further, a State has the discretion to include in its assessment system components beyond the requirements of the ESEA, which are not subject to assessment peer review. For example, some States administer assessments in additional content areas (e.g., social studies, art). A State also may include additional measures, such as formative and interim assessments, in its State assessment system.

Assessment peer review is the process through which a State documents the technical soundness of its assessment system. State success with its assessment peer review begins and hinges on the steps a State takes to develop and implement a technically sound State assessment system.

From 2005 through 2012, the Department conducted its peer review process for evaluating State assessment systems. In December 2012, in light of transitions in many States to new assessments aligned to college- and career-ready academic content standards in reading/language arts and mathematics, and advancements in the field of assessments, the Department suspended peer review of State assessment systems to review and revise the process based on current best practices in the field and lessons learned over the past decade. This 2015 revised guidance is a result of that review.

Major revisions to the assessment peer review process reflect the following:

Improvements in Educational Assessment. Numerous improvements in educational assessment have advanced the means for developing, administering, and demonstrating the technical quality of State assessments. For example, although several States have been using technology to develop and administer assessments over the past decade, the prevalence of technology is continuing to change the nature and delivery of assessments (e.g., technology-enhanced items and computer-adaptive assessments). New research regarding accessibility for students with disabilities and English learners is increasingly informing the design and development of general and alternate assessments. Similarly, advances in areas such as State test security practices and automated scoring have provided an opportunity for the Department to refine its guidance on how a State can demonstrate the quality and soundness of its assessment system. This revised guidance reflects both the expanded possibilities for State assessments and new means by which States may address the critical elements.

Revisions to Nationally Recognized Professional and Technical Standards. Section 1111(b)(3)(C)(iii) of the ESEA requires that State assessments be consistent with relevant, nationally

recognized professional and technical standards. The *Standards for Educational and Psychological Testing,*[2] nationally recognized professional and technical standards for educational assessment, were updated in Summer 2014. With a focus on the components of State assessment systems required by the ESEA, the Department's guidance reflects the revised *Standards for Educational and Psychological Testing*. This guidance both increases the emphasis on the technical quality of assessments, as reflected in section 2 of the critical elements ("Assessment System Operations"), and maintains a correspondence to the *Standards for Educational and Psychological Testing* in other areas of technical quality. As in the *Standards for Educational and Psychological Testing,* this guidance incorporates increased attention to fairness and accessibility and the expanding role of technology in developing and administering assessments.

Emergence of Multi-State Assessment Groups. Many States began administering new assessments in 2014-2015, often through participation in assessment consortia formed to develop new general and alternate assessments aligned to college- and career-ready academic content standards. This guidance responds to these developments and adapts the assessment peer review process for States participating in assessment consortia to reduce burden and to ensure consistency in the review and evaluation of State assessment systems.

Lessons Learned from the Previous Assessment Peer Review Process. The previous assessment peer review process focused on an evidence-based review by a panel of external assessment experts with technical and practical experiences with State assessment systems. This contributed to improved quality of State assessment systems. As a result, this guidance also focuses on an evidence-based review by a panel of external assessment experts.

The Department also received feedback requesting greater transparency, consistency and clarity about what is required to address the critical elements. This guidance responds by including (1) more specific examples of evidence to which States can refer in preparing their submissions and that assessment peer reviewers can use as a reference to promote consistency in their reviews, and (2) additional details about the assessment peer review process.

This guidance neither creates nor confers any rights for or on any person, nor imposes any requirements beyond those required under applicable law and regulations. This guidance represents the Department's current thinking on the critical elements and best practices for State development and implementation of assessment systems, and it supersedes the Department's previous guidance, entitled *Standards and Assessments Peer Review Guidance: Information and Examples for Meeting Requirements of the No Child Left Behind Act of 2001, revised December 21, 2007 to include modified academic achievement standards (Revised with technical edits January 12, 2009*).

[2] American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME) (2014). *Standards for Educational and Psychological Testing*. Washington DC: AERA.

B. THE ASSESSMENT PEER REVIEW PROCESS

Overview

The Department's review of State assessment systems is an evidence-based, peer review process for which each State submits evidence to demonstrate that its assessment system meets a set of established criteria, called critical elements.

Critical Elements. The critical elements in Part II of this document represent the ESEA statutory and regulatory requirements that State assessment systems must meet. The six sections of critical elements that cover these requirements are: (1) Statewide System of Standards and Assessments, (2) Assessment System Operations, (3) Technical Quality – Validity, (4) Technical Quality – Other, (5) Inclusion of All Students, and (6) Academic Achievement Standards and Reporting. The map of critical elements included in Part II provides an overview of the six sections and the critical elements within each section.

Evidence-Based Review. Each State must submit evidence for its assessment system that addresses each critical element. Consistent with sections 1111(b)(1)(A) and 1111(e)(1)(F) of the ESEA, the Department does not require a State to submit its academic content standards as part of the peer review. In addition, the Department will not require a State to include or delete any specific content in its academic content standards and a State is not required to use specific academic assessment instruments or items. The Department's assessment peer review focuses on the processes for assessment development employed by the State and the resulting evidence that confirms the technical quality of the State's assessment system.

Scheduling. The Department will notify States of the schedule for upcoming assessment peer reviews. A State implementing new assessments or a State that has made significant changes to previously reviewed assessments should submit its assessment system for assessment peer review approximately six months after the first operational administration of its new or significantly changed assessments, or the next available scheduled peer review, if applicable, and prior to the second administration of the new or revised assessment system (see also Exhibit 1).

Expert Peer Reviewers. To determine if a State has met ESEA standards and assessment requirements, the Department uses a peer review process involving experts in the field of educational standards and assessments. Based on the evidence a State submits, the reviewers evaluate the State's assessment system against ESEA requirements and provide their evaluations to the Assistant Secretary of Elementary and Secondary Education.

Selection of Assessment Peer Reviewers. Assessment peer reviewers are individuals who have strong technical expertise necessary for reviewing State assessment systems and practical experiences in applying that expertise to the operation of State assessment systems. Each assessment peer reviewer is selected by the Department based on the individual's experience and expertise, with an emphasis on knowledge of technical aspects of large-scale assessments, experience with the operation of State assessment systems, and relevant specialized expertise (such as in the relevant content areas, with technology-based assessments or with alternate assessments based on alternate academic achievement standards). Assessment peer reviewers are selected from individuals who have previously served as assessment peer reviewers for the Department or as reviewers and

consultants for other assessment-related activities for the Department; recommendations by Department staff; and recommendations from the field. Assessment peer reviewers are screened to ensure they do not have a conflict of interest.

Role of Assessment Peer Reviewers. Using the critical elements in this guidance as a framework, assessment peer reviewers apply their professional judgment and relevant professional experiences to evaluate the degree to which evidence provided about a State's assessment system addresses each of the critical elements. Their evaluations inform the decision by the Assistant Secretary of Elementary and Secondary Education as to whether or not the State has sufficiently demonstrated that its assessment system addresses each critical element.

Assessment peer reviewers work in teams to review evidence submitted by a State. Assessment peer reviewers and teams are selected by the Department to review each State's submission of evidence. The Department aims to select teams that balance peer reviewer expertise and experience in general, and, as applicable, include the expertise and experience needed for the particular assessments a State has submitted for assessment peer review (e.g., technology-based assessments or alternate assessments based on alternate academic achievement standards). The final configuration of an assessment peer review team, typically three reviewers, is determined by the Department. To protect the integrity of the assessment peer review process, the identity of the assessment peer review team for a specific State will remain anonymous.

During the peer review, the first step is for each of the assessment peer reviewers to independently review the materials submitted by a State and record their evaluation on an assessment peer review notes template.[3] Next, at an assessment peer review team meeting, the assessment peer reviewers discuss the State's submitted evidence with respect to each critical element, allowing the peer reviewers to strengthen their understanding of the evidence and to inform their individual evaluations. If there are questions or additional evidence appears to be needed, the Department may facilitate a conversation or communication between the peer reviewers and the State to clarify the State's evidence. Based upon each peer reviewer's review of the State's documentation, he or she will note where additional evidence related to or changes in a State assessment system may be necessary for the State to meet the ESEA or ESEA flexibility requirements; assessment peer reviewers may also present suggestions for addressing the outstanding requirements or highlight best practices in their notes. Although the assessment peer reviewers on a team are expected to generate one set of assessment peer review notes that reflect their review and evaluation of the State's evidence, they are not expected to reach consensus.

The assessment peer review notes serve two purposes. First, they serve as the record of the assessment peer review team's evaluation of a State's evidence for the Assistant Secretary. Second, soon after the assessment peer review ends, the assessment peer review notes are sent to the State as technical assistance and preliminary feedback prior to a formal decision regarding the outcome of the review. The assessment peer review notes, however, do not necessarily identify the final set of additional evidence, if any, that a State may need to submit to demonstrate that its assessment system meets all of the critical elements.

Training and Support for Assessment Peer Reviewers. Assessment peer reviewers will be trained in interactive training sessions with other assessment peer reviewers. Training will be based

[3] Forthcoming; will be available at: www.ed.gov/admins/lead/account/saa.html.

primarily on: (1) this guidance, and (2) the Department's Instructions for Assessment Peer Reviewers,[4] which include information on applying the critical elements to the review. Prior to the assessment peer review team meeting, each member of an assessment peer review team will be sent the following items: materials submitted to the Department by a State; this guidance; an assessment peer review notes template; and instructions for assessment peer reviewers. This allows for a thorough and independent review of the evidence based on this guidance in preparation for the assessment peer review team meeting.

Role of Department Staff. For some critical elements that serve as compliance checks or checks on processes, Department staff will review the evidence submitted by a State, as shown in the map of the critical elements. Department staff will determine either that the requirement has been adequately addressed or forward the evidence to the assessment peer review team for further review. In addition, one or more Department staff will be assigned as a liaison to each State participating in an assessment peer review and to the assessment peer review team for that State throughout the assessment peer review process. The Department liaison will serve as a contact and support for the State and assessment peer review team.

Outcomes of Assessment Peer Review. Following a submission of evidence for assessment peer review, a State first will receive feedback in the form of assessment peer review notes. Assessment peer review notes do not constitute a formal decision by the Assistant Secretary of Elementary and Secondary Education. Instead, they provide initial feedback regarding the assessment peer reviewers' evaluation and recommendations based on the evidence submitted by the State. A State should consider such feedback as technical assistance and not as formal feedback or direction to make changes to its assessment system.

The Assistant Secretary will provide formal feedback to a State regarding whether or not the State has provided sufficient evidence to demonstrate that its assessment system meets all applicable ESEA statutory and regulatory requirements following the assessment peer review. If a State has not provided sufficient evidence, the Assistant Secretary will identify the additional evidence necessary to address the critical elements. The Department will work with the State to develop a plan and timeline for submitting the additional evidence for assessment peer review.

Assessment Peer Review and Civil Rights Compliance. The assessment peer review will not evaluate or provide recommendations regarding whether or not a State's assessment system complies with Federal civil rights laws that prohibit discrimination based on race, color, national origin, sex, disability, and age. These laws include Title VI of the Civil Rights Act of 1964, Title IX of the Education Amendments of 1972, Section 504 of the Rehabilitation Act of 1973, Title II of the Americans with Disabilities Act, the Age Discrimination Act of 1975, and applicable requirements in the Individuals with Disabilities Education Act.

[4] Forthcoming; will be available at: www.ed.gov/admins/lead/account/saa.html.

Requirements for Assessment Peer Review When a State Makes a Change to a Previously Peer Reviewed State Assessment System

In general, a significant change to a State assessment system is one that changes the interpretation of test scores. If a State makes a significant change to a component of its State assessment system that the State has previously submitted for assessment peer review, the State must submit evidence related to the affected component for assessment peer review. A State should submit evidence for assessment peer review before, or as soon as reasonable following, the first administration of its assessment system with the change and no later than prior to the second administration of its assessment system with the change.

To provide clarity about the implications of changes to State assessment systems, this guidance outlines three categories of changes (see Exhibit 1 for further details):

- *Significant change* clearly changes the interpretation of test scores and requires a new assessment peer review.
- *Adjustment* is a change between the extremes of significant and inconsequential and may or may not require a new assessment peer review.
- *Inconsequential change* is a minor change that does not impact the interpretation of test scores or substantially change other key aspects of a State's assessment system. For inconsequential changes, a new assessment peer review is not required.

A State making a change to its assessment system is encouraged to discuss the implications of the change for assessment peer review with its technical advisory committee (TAC). A State making a significant change or adjustment to its assessment system also is encouraged to contact the Department early in the planning process to determine if the adjustment is significant and to develop an appropriate timeline for the State to submit evidence related to significant changes for assessment peer review. Exhibit 1 provides examples of the three categories of changes. However, as noted in the exhibit, the changes listed in Exhibit 1 are merely illustrative and do not constitute an exhaustive list of the changes that fall within each category.

Exhibit 1: Categories of Changes and Non-Exhaustive Examples of Assessment Peer Review Submission Requirements when a State Makes a Change to a Previously Peer Reviewed State Assessment System

New Assessments

Submission must address: Sections 2, 3, 4, 5 and 6 of the critical elements

Significant	Always significant.
Adjustment	Not applicable.
Inconsequential	Not applicable.

Development of a Technology-Based Version of an Assessment

Submission must address: Critical elements 2.1–2.3, sections 3 and 4 of the critical elements

Significant	• Assessment delivery is changed from entirely paper-and-pencil to entirely computer-based. • The new computer-based version of the assessment includes technology-enhanced items that are not available in the simultaneously administered paper-and-pencil version.
Adjustment	• Assessment delivery is changed from a mix of paper-and-pencil and computer-based assessments to an entirely technology-based administration using a range of devices.
Inconsequential	• Not applicable.

Development of a Native Language Version of an Assessment

Submission must address: Critical elements 2.1 - 2.3, sections 3 and 4 of the critical elements

Significant	Always significant.
Adjustment	Not applicable.
Inconsequential	Not applicable.

Changes to an Existing Test Design

Submission must address: Case specific; likely from sections 2, 3, 4 and 6 of the critical elements

Significant	• State's approved reading/language arts assessment included multiple constructed-response items for each academic content standard and the State replaces these items with multiple-choice items. • Change in the assessment purpose, use, design, or content (e.g., due to legislative or regulatory change or State policy).
Adjustment	• State changes the number of items on its assessments (by more than a few items). • Scoring conducted across an assessment consortium as a whole is changed to scoring conducted individually by consortium member States.
Inconsequential	• State changes from on-site to virtual training for its scorers of extended-response items.

Changes to Test Administration

Submission must address: Case specific; likely from sections 2 and 5 of the critical elements

Significant	• State shifts scoring of its alternate assessments based on alternate achievement standards from centralized third-party scoring to scoring by the test administrator.
Adjustment	• State shifts from desktop computer-based test administration to Statewide test administration using a range of devices. • State participates in an assessment consortium and shifts certain practices from consortium-level to State-level operation.
Inconsequential	• State combines its trainings for test security and test administration.

Assessments Based on New Academic Achievement Standards

Submission must address: Section 6 of the critical elements

Significant	• Comprehensive revision of State's academic achievement standards (e.g., performance-level descriptors, cut-scores).
Adjustment	• Smoothing of cut-score across grades after multiple administrations.
Inconsequential	• Implementation of planned adjustment to State's achievement standards that were reviewed and approved during assessment peer review.

Assessments Based on New or Revised Academic Content Standards

Submission must address: Sections 1, 2, 3, 4, 5 and 6 of the critical elements

Significant	• Adoption of completely new academic content standards or comprehensive revision of the State's academic content standards to which assessments must be aligned.
Adjustment	• State makes minor changes to its academic content standards to which assessments must be aligned by moving a few benchmarks within or between standards, but assessment blueprints are not affected and test results are comparable after equating.
Inconsequential	• State makes minor formatting changes, non-substantive word edits or corrections to typos in its academic content standards, which do not change the interpretation of test scores.

C. PREPARING AN ASSESSMENT PEER REVIEW SUBMISSION

A State should use this guidance to prepare its submission for assessment peer review. The State Assessment Peer Review Submission Index Template includes a checklist a State can use to prepare an assessment peer review submission.

Content and Organization of a State Submission for Assessment Peer Review

Submission by State. A State should send its submission to the Department according to the schedule for assessment peer reviews announced by the Department. A State should submit its assessment systems for assessment peer review approximately six months after the first operational administration of new or significantly changed assessments. The Department encourages each State to plan for preparing its peer review submission according to this timeline. A State is expected to submit evidence regarding its State assessment system approximately three weeks prior to the scheduled assessment peer review date for the State.

For assessments administered by multiple States, the Department will conduct a single review of the evidence that applies to all States implementing the same assessments. This approach both promotes consistency in the review of such assessments and reduces burden on States in preparing for assessment peer review.

What to Include in a Submission. A State's submission should include the following parts:

1) State Assessment Peer Review Submission Cover Sheet;
2) State Assessment Peer Review Submission Index, as described below; and
3) Evidence to address each critical element.

State Assessment Peer Review Submission Cover Sheet and Index Template.[5] The *State Assessment Peer Review Submission Cover Sheet and Index Template* includes the cover sheet that a State must submit with each submission of evidence for peer review. It also includes an index template aligned to the critical elements in six sections as shown in the map of the critical elements. Each State should use this State Assessment Peer Review Submission Index Template to prepare an index to its submission to accompany the evidence that the State submits. A State's prepared index should outline the evidence for each critical element with the following:

1) Identification of the critical element;
2) List of the evidence submitted to address the critical element (e.g., relevant document(s) and page number(s);
3) Indication of where evidence that addresses the applicable critical element can be found in the State's submission; and
4) As applicable, a brief narrative of how the evidence addresses the critical element or any explanatory notes relevant to the evidence.

Because a State will submit numerous pieces of evidence, the Department recommends that the State use a coding scheme to identify the various pieces of evidence cited in its State Assessment Peer Review Submission Index. Exhibit 3, at the end of this part of this guidance, shows suggested formats for how a State might present its submission.

[5] Available at: www.ed.gov/admins/lead/account/saa.html.

Preparing Evidence. The Department encourages a State to take into account the following considerations in preparing its submission of evidence:

The description and examples of evidence apply to each assessment in the State's assessment system (e.g., general and alternate assessments, assessments in each content area). For example, for Critical Element 2.3 – Test Administration, a State must address the critical element for both its general assessments and its alternate assessments. In general, evidence submitted should be based on the most recent year of test administration in the State.

Multiple critical elements are likely to be addressed by the same documents. In such cases, a State is encouraged to streamline its submission by submitting one copy of such evidence and cross-referencing the evidence across critical elements in the completed State Assessment Peer Review Submission Index for its submission. For example, it is likely that the test coordinator and test administration manuals, the accommodations manual, technical reports for the assessments, results of an independent alignment study, and the academic achievement standards-setting report will address multiple critical elements for a State's assessment system.

Similarly, if certain pieces of evidence are substantially the same across assessments, a sample, rather than the full set of such evidence, may be submitted. For example, if the State has submitted all of its grades 3-8 and high school reading/language arts and mathematics assessments for assessment peer review, sample individual student reports must be submitted for both general and alternate assessments under Critical Element 6.4 – Reporting. However, if the individual student reports are substantially the same across grades, the State may choose to submit a sample of the reports, such as individual student reports for both subjects for grades 3, 7, and high school.

A State should send its submission in electronic format and the files should be clearly indexed, with corresponding electronic folders, folder names, and filenames. For evidence that is typically presented in an Internet-based format, screenshots may be submitted as evidence. Links to websites should not be submitted as evidence.

Coordination of Submissions for States that Administer the Same Assessments

In the case of multiple States administering the same assessment(s), the Department will hold one assessment peer review for those assessments in order to reduce the burden on States and to promote consistency in the assessment peer review. This includes groups of States that formed consortia for the purpose of developing assessments and States that administer the same commercially developed assessments (e.g., multiple States that are all administering the same commercially developed test as their high school assessment).

For evidence that is common across an assessment administered in multiple States, the submission of evidence should be coordinated, with one State submitting the evidence on behalf of all States administering the assessment. Each State also must submit State-specific evidence that is not common among States that administer the same assessment(s). As described below, in their State-specific submissions, individual States should cross-reference coordinated submissions. A State for which a coordinated submission of evidence is part of its evidence for assessment peer review is encouraged to submit its State-specific evidence at the same time as the coordinated submission.

A specific State submitting on behalf of itself and other States that administer the same assessment(s) must identify the States on whose behalf the evidence is submitted. Correspondingly, each State administering the same assessment should include in its State-specific submission a letter that affirms that the submitting State is submitting assessment peer review evidence on its behalf.

Exhibit 2 below outlines which critical elements the Department anticipates may be addressed by evidence that is State-specific and evidence that is common among States that administer the same assessment(s). The evidence needed to fully address some critical elements may be a hybrid of the two types of evidence. For example, under Critical Element 2.3 – Test Administration, test administration and training materials may be the same across States administering the same assessment(s) while each individual State may conduct various trainings for test administration. In such an instance, the submitting State would submit the test administration and training materials, and each State would separately submit evidence regarding implementation of the actual training. This information is also displayed graphically on the map of the critical elements.

Exhibit 2: Evidence for Critical Elements that Likely Will Be Addressed by Submissions of Evidence that are State-Specific, Coordinated for States Administering the Same Assessments, or a Hybrid

Evidence	Critical Elements
State-specific evidence	1.1, 1.2, 1.3, 1.4, 1.5, 2.4, 5.1, 5.2 and 6.1
Coordinated evidence for States administering the same assessments	2.1, 2.2, 3.1, 3.2, 3.3, 3.4, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 6.2 and 6.3
Hybrid evidence	2.3, 2.5, 2.6, 5.3, 5.4 and 6.4

A State that administers an assessment that is the same as an assessment administered in other States in some ways but that also differs from the other States' assessment in certain ways should consult Department staff for technical assistance on how requirements for submission apply to the State's circumstances.

How to Read the Critical Elements

Critical Elements and Examples of Evidence. Each critical element includes two parts: the critical element and examples of evidence.

Critical Element. The critical element is a statement of the relevant requirement, and a State must submit evidence to document that its assessment system meets the requirement. The set of evidence submitted for each critical element, collectively, should address the entirety of the critical element.

Examples of Evidence. Examples of evidence associated with each critical element within Part II of this guidance are generally illustrative. A State may address the critical elements in a range of ways. The examples of evidence provided are intended to facilitate preparation of a State's assessment peer review submission by illustrating or suggesting documentation often available to States that likely would address the critical element in whole or in part. Not all of the listed evidence may be necessary for each State submission, and a State may determine that other types of evidence better address a critical element than those included in Part II of this guidance.

For technology-based assessments, some evidence, in addition to that required for other assessment modes, is needed to address certain critical elements due to the nature of technology-based assessments. In such cases, examples of evidence unique to technology-based assessments are identified with an icon of a computer.

For alternate assessments based on alternate academic achievement standards for students with the most significant cognitive disabilities (AA-AAAS), the evidence needed to address some critical elements may vary from the evidence needed for a State's general assessments due to the nature of the alternate assessments. For some critical elements, different examples of evidence are provided for AA-AAAS. For other critical elements, additional evidence to address the critical elements for AA-AAAS is listed. In such cases, examples of evidence unique to AA-AAAS are identified with an icon of "AA-AAAS."

Terminology. The following explanations of terms apply to the critical elements and examples of evidence in Part II.

Accessibility tools and features. This refers to adjustments to an assessment that are available for all test takers and are embedded within an assessment to remove construct irrelevant barriers to a student's demonstration of knowledge and skills. In some testing programs, sets of accessibility tools and features have specific labels (e.g., "universal tools" and "accessibility features").

Accommodations. For purposes of this guidance, "accommodations" generally refers to adjustments to an assessment that provide better access for a particular test taker to the assessment and do not alter the assessed construct. These are applied to the presentation, response, setting, and/or timing/scheduling of an assessment for particular test takers. They may be embedded within an assessment or applied after the assessment is designed. In some testing programs, certain adjustments may not be labeled accommodations but are considered accommodations for purposes of peer review because they are allowed only when selected for an individual student. For students eligible under IDEA or covered by Section 504, accommodations provided during assessments must be determined in accordance with 34 CFR §200.6(a) of the Title I regulations.

All public elementary and secondary schools. This includes general public schools; public charter schools; public virtual schools; and special purpose schools, such as detention and residential centers under the authority of the State educational agency; and schools that serve students with special needs (e.g., special education centers).

All public elementary and secondary school students. This includes *all* students enrolled in public schools, including English learners; students with disabilities; migratory students; students experiencing homelessness; and students placed in private schools using public funds. A student placed in a private school by a public agency for the purpose of receiving special education and related services must be included in the State assessment system.

Alternate academic achievement standards. "Alternate academic achievement standards" set expectations of performance that differ in complexity from grade-level achievement standards. A State may adopt alternate academic achievement standards for students with the most significant cognitive disabilities. Alternate academic achievement standards must: (1) be 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 for students with the most significant cognitive disabilities.

Alternate assessments. Under ESEA regulations, a State's assessment system must provide for one or more alternate assessments for students with disabilities whose IEP Teams determine that they cannot participate in all or part of the State assessments, even with appropriate accommodations. A State may administer an alternate assessment aligned to either grade-level academic achievement standards, or for students with the most significant cognitive disabilities, alternate academic achievement standards.

Alternate assessments based on alternate academic achievement standards (AA-AAAS). AA-AAAS are for use only for students with the most significant cognitive disabilities. AA-AAAS must be aligned to the State's grade-level academic content standards, but may assess the grade-level academic content standards with reduced breadth and cognitive complexity than general assessments. For an AA-AAAS, extended academic content standards often are used to show the relationship between the State's grade-level academic content standards and the content assessed on the AA-AAAS. AA-AAAS include content that is challenging for students with the most significant cognitive disabilities and may not contain unrelated content (e.g., functional skills). Evidence of how breadth and cognitive complexity are determined and operationalized should be submitted as part of Critical Element 2.1 – Test Design and Development.

Assessments aligned to the full range of a State's academic content standards. A State's assessment system under ESEA Title I must assess the depth and breadth of the State's grade-level academic content standards — i.e., be aligned to the full range of those standards. Assessing the full range of a State's academic content standards means that each State assessment covers the domains or major components within a content area. For example, if a State's academic content standards for reading/language arts identify the domains of reading, language arts, writing, and speaking and listening, assessing the full range of reading/language arts standards means that the assessment is aligned to all four of these domains. Assessing the full range of a State's standards also means that specific content in a State's academic content standards is not systematically excluded from a State's assessment system. Assessing the full range of standards, however, does not mean that each State assessment must annually cover all discrete knowledge and skills represented within a State's academic content standards; rather, assessing the full range of a State's academic standards means that a State's assessment system covers all of the knowledge and skills over a period of time. Both Critical Element 2.1 – Test Design and Development and Critical Element 3.1 – Overall Validity, including Validity Based on Content examine whether a State's assessment system is aligned to the full range of the State's academic content standards.

In addition to ensuring that each assessment covers the full range of the domains or major components represented in a State's grade-level academic content standards for a content area, a State may include additional content from adjacent grades in its assessments to provide additional information to parents and teachers regarding student achievement. If a State includes content for both the grade in which a student is enrolled and adjacent grades, assessing the full range of a State's academic content standards means: (1) that each State assessment assesses the full range of the State's content standards for the tested grade, as described above, (2) that the assessment provides a score for the student that is based only on the student's performance on grade-level academic content standards, and (3) that each student's score is at least as precise as the score for a student assessed only on grade-level academic content standards. Because assessing off-grade-level content

is, by definition, not part of assessing the full range of a State's grade-level academic content standards, evidence for assessment peer review (e.g., validity studies, achievement standards-setting reports) that reflects the inclusion of off-grade-level content would not be applicable to addressing the critical elements, and only student performance based on grade-level academic content and achievement standards would meet accountability and reporting requirements under Title I.

Collectively. For some critical elements, the expectation is that a body of evidence will be required and the sum of the various pieces of evidence will be considered for the critical element. For example, for Critical Element 3.1 – Overall Validity, including Validity Based on Content, the evidence will be evaluated to determine if, collectively, it presents sufficient validity evidence for the State's assessments.

For such critical elements, the State should provide a summary of the body of evidence addressing the critical element, in addition to the individual pieces of evidence. Ideally, this summary is something the State has prepared for its own use (e.g., a chapter in the technical report for its assessments or a report to its TAC), as opposed to a summary created solely for the State's assessment peer review submission. Such critical elements are indicated by beginning the corresponding descriptions of examples of evidence with the word, "collectively."

Evidence. Evidence means documentation related to a State's assessment system that is used to address a critical element, such as State statutes or regulations; technical reports; test coordinator and administration manuals; and summaries of analyses. As much as possible, a State should rely for evidence on documentation created in the development and operation of its assessment system, in contrast to documentation prepared primarily for assessment peer review.

In general, the examples of evidence for critical elements refer to two types of evidence: procedural evidence and empirical evidence. Procedural evidence generally refers to steps taken by a State in developing and administering the assessment, and empirical evidence generally refers to analyses that confirm the technical quality of the assessments. For example, Critical Element 4.2 – Fairness and Accessibility requires procedural evidence that the assessments were designed and developed to be fair and empirical evidence that confirms they were fair when actually administered.

Key documents. Submitted evidence should reflect the State's assessment system and the State's standard, routine procedures for implementing its assessments. In addition, such assessment materials for districts and schools (e.g., test coordinator manuals, test administration manuals, accommodations manuals, etc.) should be consistent across assessments included in the State's assessment system. To indicate cases in which it is especially important for "key" documents to be submitted as evidence, the term "key" is used in the critical element or examples of evidence.

Exhibit 3: Examples of a Prepared State Index for Selected Critical Elements

Critical Element 4.2 – Fairness and Accessibility (EXAMPLE)

	Evidence	Notes
The State has taken reasonable and appropriate steps to ensure that its assessments are accessible to all students and fair across student groups in the design, development and analysis of its assessments.	**General assessments in reading/language arts and mathematics:** Evidence #24: Technical Manual (2015). The technical manual for the State assessments documents steps taken to ensure fairness: • Pp. 30-37 discuss steps taken during design and development. • Pp. 86-92 discuss analyses of assessment data. Evidence #25: Summary of follow-up to differential item functioning (DIF) analysis. Evidence #26: Amendment to assessment contract requiring additional bias review for items and added instructions for future item development. **Alternate assessments in reading/language arts and mathematics:** The State's alternate assessments were developed by the ABC assessment consortium. Evidence for the assessments was submitted on this State's behalf by State X. (See State Assessment Peer Review Submission Cover Sheet)	**General assessments in reading/language arts and mathematics:** • DIF analyses showed differences by gender for several items in reading/language arts assessments for the grades 3 and 4. Examination of the items showed they all involved reading informational text. To address this for the next test administration, a sensitivity review of all grade 3 and 4 reading/language passages involving informational text will undergo an additional bias review. Instructions for item development in future years will be revised to address this as well. **Alternate assessments in reading/language arts and mathematics:** No notes.

Why this works:

- Concise and clearly written
- Evidence, including page numbers, clearly identified
- Content areas addressed and clearly identified
- Both general and alternate assessments addressed, as appropriate
- Where evidence identified shortcoming, notes discuss how State is addressing
- Cross references submission for assessment consortium

Critical Element 5.2 – Procedures for Including English Learners (EXAMPLE)

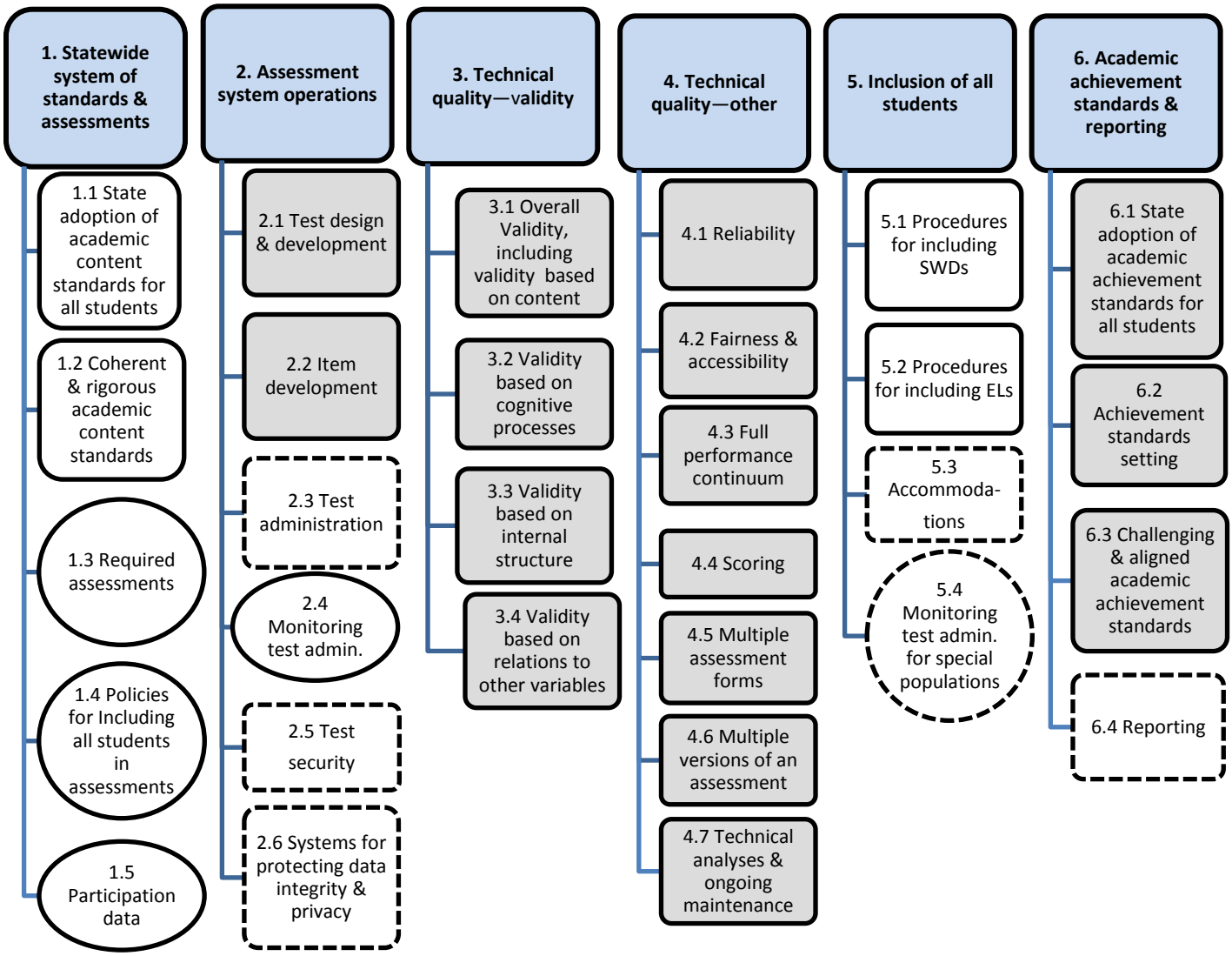
	Evidence	**Notes**
The State has in place procedures to ensure the inclusion of all English learners in public schools in the State's assessment system and clearly communicates this information to districts, schools, teachers, and parents including, at a minimum: • Procedures for determining whether an English learner should be assessed with accommodation(s); • Information on accessibility tools and features available to all students and assessment accommodations available for English learners; • Guidance regarding selection of appropriate accommodations for English learners.	The State's procedures for determining whether an English learner should be assessed with accommodation(s) on either the State's general assessment or AA-AAAS are in: - Instructions for Student Language Acquisition Plans for English Learners; - Template for Language Acquisition Plan for English Learners. For the general assessments, information on accessibility tools and features is in: - District Test Coordinator Manual (see p. 5) - School Test Coordinator Manual (see p. 7) - Test administrator manuals (grade 3 for reading/language arts, grade 8 for math – see p. 5 of each) For the general assessments, guidance regarding selection of accommodations is in: - State Accommodations Manual (see pp. 23-32). For the AA-AAAS, information on accessibility tools and features and accommodations is in: - District Test Coordinator Manual (see p. 6) - School Test Coordinator Manual (see p. 5) - Test administrator manuals (see p. 5) Evidence: - Folder 6, File #22 – Instructions for student Language Acquisition Plan for English Learners - Folder 6, File #23 – Template for Language Acquisition Plan for English Learners. - Folder 6, File #24 - District Test Coordinator Manual; - Folder 6, File #25 - School Test Coordinator Manual; - Folder 6, File #26 – Grade 3 Reading/language Arts Test Administration Manual - Folder 7, File #27 – Grade 8 Math Test Administration Manual - Folder 8, File #28 – Grade 10 Reading/language Arts and Math Administration Manual) - Folder 9, File #29 -- State Accommodations Manual	The State's Language Acquisition Plan for English Learners applies to both students who take general assessments and students who take the State's AA-AAAS.

Why this works:

- Concise and clearly written
- Evidence, including page numbers, clearly identified
- Addresses coordination and consistency across assessments, as appropriate
- Notes provided only where helpful

II – CRITICAL ELEMENTS FOR ASSESSMENT PEER REVIEW

Map of the Critical Elements for the State Assessment System Peer Review



KEY

- Critical elements in ovals will be checked for completeness by Department staff; if necessary, they may also be reviewed by assessment peer reviewers (e.g., Critical Element 1.3). All other critical elements will be reviewed by assessment peer reviewers.
- Critical elements in shaded boxes likely will be addressed by coordinated evidence for all States administering the same assessments (e.g., Critical Element 2.1).
- Critical elements in clear boxes with solid outlines likely will be addressed with State-specific evidence, even if a State administers the same assessments administered by other States (e.g., Critical Element 5.1).
- Critical elements in ovals or clear boxes with dashed outlines likely will be addressed by both State-specific evidence and coordinated evidence for States administering the same assessments (e.g., Critical Element 2.3, 5.4).

SECTION 1: STATEWIDE SYSTEM OF STANDARDS AND ASSESSMENTS

Critical Element 1.1 – State Adoption of Academic Content Standards for All Students

	Examples of Evidence

The State formally adopted challenging academic content standards for all students in reading/language arts, mathematics and science and applies its academic content standards to all public elementary and secondary schools and students in the State.	Evidence to support this critical element for the State's assessment system includes:
	• Evidence of adoption of the State's academic content standards, specifically:
	  o Indication of *Requirement Previously Met*; or
	  o State Board of Education minutes, memo announcing formal approval from the Chief State School Officer to districts, legislation, regulations, or other binding approval of a particular set of academic content standards;
	• Documentation, such as text prefacing the State's academic content standards, policy memos, State newsletters to districts, or other key documents, that explicitly state that the State's academic content standards apply to all public elementary and secondary schools and all public elementary and secondary school students in the State;
	
	Note: A State with *Requirement Previously Met* should note the applicable category in the State Assessment Peer Review Submission Index for its peer submission. *Requirement Previously Met* applies to a State in the following categories: (1) a State that has academic content standards in reading/language arts, mathematics, or science that have not changed significantly since the State's previous assessment peer review; or (2) with respect to academic content standards in reading/language arts and mathematics, a State approved for ESEA flexibility that (a) has adopted a set of college- and career-ready academic content standards that are common to a significant number of States and has not adopted supplemental State-specific academic content standards in these content areas, or (b) has adopted a set of college- and career-ready academic content standards certified by a State network of institutions of higher education (IHEs).

Critical Element 1.2 – Coherent and Rigorous Academic Content Standards

	Examples of Evidence

The State's academic content standards in reading/language arts, mathematics and science specify what students are expected to know and be able to do by the time they graduate from high school to succeed in college and the workforce; contain content that is coherent (e.g., within and across grades) and rigorous; encourage the teaching of advanced skills; and were developed with broad stakeholder involvement.	Evidence to support this critical element for the State's assessment system includes:
	• Indication of *Requirement Previously Met;* or
	• Evidence that the State's academic content standards:
	  o Contain coherent and rigorous content and encourage the teaching of advanced skills, such as:
	  ▪ A detailed description of the s**t**rategies the State used to ensure that its academic content standards adequately specify what students should know and be able to do;
	  ▪ Documentation of the process used by the State to benchmark its academic content standards to nationally or internationally recognized academic content standards;
	  ▪ Reports of external independent reviews of the State's academic content standards by content experts, summaries of reviews by educators in the State, or other documentation to confirm that the State's academic content standards adequately specify what students should know and be able to do;

	▪ Endorsements or certifications by the State's network of institutions of higher education (IHEs), professional associations and/or the business community that the State's academic content standards represent the knowledge and skills in the content area(s) under review necessary for students to succeed in college and the workforce; o Were developed with broad stakeholder involvement, such as: ▪ Summary report of substantive involvement and input of educators, such as committees of curriculum, instruction, and content specialists, teachers and others, in the development of the State's academic content standards; ▪ Documentation of substantial involvement of subject-matter experts, including teachers, in the development of the State's academic content standards; ▪ Descriptions that demonstrate a broad range of stakeholders was involved in the development of the State's academic content standards, including individuals representing groups such as students with disabilities, English learners and other student populations in the State; parents; and the business community; ▪ Documentation of public hearings, public comment periods, public review, or other activities that show broad stakeholder involvement in the development or adoption of the State's academic content standards. Note: See note in Critical Element 1.1 – State Adoption of Academic Content Standards for All Students. With respect to academic content standards in reading/language arts and mathematics, *Requirement Previously Met* does not apply to supplemental State-specific academic content standards for a State approved for ESEA flexibility that has adopted a set of college- and career-ready academic content standards in a content area that are common to a significant number of States and also adopted supplemental State-specific academic content standards in that content area.

Critical Element 1.3 – Required Assessments

	Examples of Evidence
The State's assessment system includes annual general and alternate assessments (based on grade-level academic achievement standards or alternate academic achievement standards) in: • Reading/language arts and mathematics in each of grades 3-8 and at least once in high school (grades 10-12); • Science at least once in each of three grade spans (3-5, 6-9 and 10-12).	Evidence to support this critical element for the State's assessment system includes: • A list of the annual assessments the State administers in reading/language arts, mathematics and science including, as applicable, alternate assessments based on grade-level academic achievement standards or alternate academic achievement standards for students with the most significant cognitive disabilities, and native language assessments, and the grades in which each type of assessment is administered.

Critical Element 1.4 – Policies for Including All Students in Assessments

	Examples of Evidence
The State requires the inclusion of all public elementary and secondary school students in its assessment system and clearly and consistently communicates this requirement to districts and schools. • For students with disabilities, policies state that all students with disabilities in the State, including students with disabilities publicly placed in private schools as a means of providing special education and related services, must be included in the assessment system; • For English learners: o Policies state that all English learners must be included in the assessment system, unless the State exempts a student who has attended schools in the U.S. for less than 12 months from one administration of its reading/	Evidence to support this critical element for the State's assessment system includes documents such as: • Key documents, such as regulations, policies, procedures, test coordinator manuals, test administrator manuals and accommodations manuals that the State disseminates to educators (districts, schools and teachers), that clearly state that all students must be included in the State's assessment system and do not exclude any student group or subset of a student group; • For students with disabilities, if needed to supplement the above: Instructions for Individualized Education Program (IEP) teams and/or other key documents; • For English learners, if applicable and needed to supplement the above: Test administrator manuals and/or other key documents that show that the State provides a native language (e.g., Spanish, Vietnamese) version of its assessments.

language arts assessment; o If the State administers native language assessments, the State requires English learners to be assessed in reading/language arts in English if they have been enrolled in U.S. schools for three or more consecutive years, except if a district determines, on a case-by-case basis, that native language assessments would yield more accurate and reliable information, the district may assess a student with native language assessments for a period not to exceed two additional consecutive years.	

Critical Element 1.5 – Participation Data

	Examples of Evidence
The State's participation data show that all students, disaggregated by student group and assessment type, are included in the State's assessment system. In addition, if the State administers end-of-course assessments for high school students, the State has procedures in place for ensuring that each student is tested and counted in the calculation of participation rates on each required assessment and provides the corresponding data.	Evidence to support this critical element for the State's assessment system includes: • Participation data from the most recent year of test administration in the State, such as in Table 1 below, that show that all students, disaggregated by student group (i.e., students with disabilities, English learners, economically disadvantaged students, students in major racial/ethnic categories, migratory students, and male/female students) and assessment type (i.e., general and AA-AAAS) in the tested grades are included in the State's assessments for reading/language arts, mathematics and science; • If the State administers end-of-course assessments for high school students, evidence that the State has procedures in place for ensuring that each student is included in the assessment system during high school, including students with the most significant cognitive disabilities who take an alternate assessment based on alternate academic achievement standards and recently arrived English learners who take an ELP assessment in lieu of a reading/language arts assessment, such as: o Description of the method used for ensuring that each student is tested and counted in the calculation of participation rate on each required assessment. If course enrollment or another proxy is used to count all students, a description of the method used to ensure that all students are counted in the proxy measure; o Data that reflect implementation of participation rate calculations that ensure that each student is tested and counted for each required assessment. Also, if course enrollment or another proxy is used to count all students, data that document that all students are counted in the proxy measure.

Table 1: Students Tested by Student Group in [subject] during [school year]

Student group		Gr 3	Gr 4	Gr 5	Gr 6	Gr 7	Gr 8	HS
All	# enrolled							
	# tested							
	% tested							
Economically disadvantaged	# enrolled							
	# tested							
	% tested							
Students with disabilities	# enrolled							
	# tested							
	% tested							
(Continued for all other student groups)	# enrolled							
	# tested							
	% tested							

Number of students assessed on the State's AA-AAAS in [subject] during [school year]: _____.

Note: A student with a disability should only be counted as tested if the student received a valid score on the State's general or alternate assessments submitted for assessment peer review for the grade in which the student was enrolled. If the State permits a recently arrived English learner (i.e., an English learner who has attended schools in the U.S. for less than 12 months) to be exempt from one administration of the State's reading/language arts assessment and to take the State's ELP assessment in lieu of the State's reading/language arts assessment, then the State should count such students as 'tested' in data submitted to address critical element.

SECTION 2: ASSESSMENT SYSTEM OPERATIONS

Critical Element 2.1 – Test Design and Development

	Examples of Evidence
The State's test design and test development process is well-suited for the content, is technically sound, aligns the assessments to the full range of the State's academic content standards, and includes: • Statement(s) of the purposes of the assessments and the intended interpretations and uses of results; • Test blueprints that describe the structure of each assessment in sufficient detail to support the development of assessments that are technically sound, measure the full range of the State's grade-level academic content standards, and support the intended interpretations and uses of the results; • Processes to ensure that each assessment is tailored to the knowledge and skills included in the State's academic content standards, reflects appropriate inclusion of challenging content, and requires complex demonstrations or applications of knowledge and skills (i.e., higher-order thinking skills); • If the State administers computer-adaptive assessments, the item pool and item selection procedures adequately support the test design.	Evidence to support this critical element for the State's general assessments and AA-AAAS includes: For the State's general assessments: • Relevant sections of State code or regulations, language from contract(s) for the State's assessments, test coordinator or test administrator manuals, or other relevant documentation that states the purposes of the assessments and the intended interpretations and uses of results; • Test blueprints that: o Describe the structure of each assessment in sufficient detail to support the development of a technically sound assessment, for example, in terms of the number of items, item types, the proportion of item types, response formats, range of item difficulties, types of scoring procedures, and applicable time limits; o Align to the State's grade-level academic content standards in terms of content (i.e. knowledge and cognitive process), the full range of the State's grade-level academic content standards, balance of content, and cognitive complexity; • Documentation that the test design that is tailored to the specific knowledge and skills in the State's academic content standards (e.g., includes extended response items that require demonstration of writing skills if the State's reading/language arts academic content standards include writing); • Documentation of the approaches the State uses to include challenging content and complex demonstrations or applications of knowledge and skills (i.e., items that assess higher-order thinking skills, such as item types appropriate to the content that require synthesizing and evaluating information and analytical text-based writing or multiple steps and student explanations of their work); for example, this could include test specifications or test blueprints that require a certain portion of the total score be based on item types that require complex demonstrations or applications of knowledge and skills and the rationale for that design.  For the State's technology-based general assessments, in addition to the above: • Evidence of the usability of the technology-based presentation of the assessments, including the usability of accessibility tools and features (e.g., embedded in test items or available as an accompaniment to the items), such as descriptions of conformance with established accessibility standards and best practices and usability studies; • For computer-adaptive general assessments: o Evidence regarding the item pool, including: ▪ Evidence regarding the size of the item pool and the characteristics (non-statistical (e.g., content) and statistical) of the items it contains that demonstrates that the item pool has the capacity to produce test forms that adequately reflect the State's test blueprints in terms of: - Full range of the State's academic content standards, balance of content, cognitive complexity for

	each academic content standard, and range of item difficulty levels for each academic content standard; - Structure of the assessment (e.g., numbers of items, proportion of item types and response types); o Technical documentation for item selection procedures that includes descriptive evidence and empirical evidence (e.g., simulation results that reflect variables such as a wide range of student behaviors and abilities and test administration early and late in the testing window) that show that the item selection procedures are designed adequately for: ▪ Content considerations to ensure test forms that adequately reflect the State's academic content standards in terms of the full range of the State's grade-level academic content standards, balance of content, and the cognitive complexity for each standard tested; ▪ Structure of the assessment specified by the blueprints; ▪ Reliability considerations such that the test forms produce adequately precise estimates of student achievement for all students (e.g., for students with consistent and inconsistent testing behaviors, high- and low-achieving students; English learners and students with disabilities); ▪ Routing students appropriately to the next item or stage; ▪ Other operational considerations, including starting rules (i.e., selection of first item), stopping rules, and rules to limit item over-exposure. **AA-AAAS.** For the State's AA-AAAS: • Relevant sections of State code or regulations, language from contract(s) for the State's assessments, test coordinator or test administrator manuals, or other relevant documentation that states the purposes of the assessments and the intended interpretations and uses of results for students tested; • Description of the structure of the assessment, for example, in terms of the number of items, item types, the proportion of item types, response formats, types of scoring procedures, and applicable time limits. For a portfolio assessment, the description should include the purpose and design of the portfolio, exemplars, artifacts, and scoring rubrics; • Test blueprints (or, where applicable, specifications for the design of portfolio assessments) that reflect content linked to the State's grade-level academic content standards and the intended breadth and cognitive complexity of the assessments; • To the extent the assessments are designed to cover a narrower range of content than the State's general assessments and differ in cognitive complexity: o Description of the breadth of the grade-level academic content standards the assessments are designed to measure, such as an evidence-based rationale for the reduced breadth within each grade and/or comparison of intended content compared to grade-level academic content standards; o Description of the strategies the State used to ensure that the cognitive complexity of the assessments is appropriately challenging for students with the most significant cognitive disabilities; o Description of how linkage to different content across grades/grade spans and vertical articulation of academic expectations for students is maintained; • If the State developed extended academic content standards to show the relationship between the State's grade-level academic content standards and the content of the assessments, documentation of their use in the design

	of the assessments; • For adaptive alternate assessments (both computer-delivered and human-delivered), evidence, such as a technical report for the assessments, showing: o Evidence that the size of the item pool and the characteristics of the items it contains are appropriate for the test design; o Evidence that rules in place for routing students are designed to produce test forms that adequately reflect the blueprints and produce adequately precise estimates of student achievement for classifying students; o Evidence that the rules for routing students, including starting (e.g., selection of first item) and stopping rules, are appropriate and based on adequately precise estimates of student responses, and are not primarily based on the effects of a student's disability, including idiosyncratic knowledge patterns; • For technology-based AA-AAAS, in addition to the above, evidence of the usability of the technology-based presentation of the assessments, including the usability of accessibility tools and features (e.g., embedded in test items or available as an accompaniment to the items), such as descriptions of conformance with established accessibility standards and best practices and usability studies.

Critical Element 2.2 – Item Development

	Examples of Evidence
The State uses reasonable and technically sound procedures to develop and select items to assess student achievement based on the State's academic content standards in terms of content and cognitive process, including higher-order thinking skills.	Evidence to support this critical element for the State's general assessments and AA-AAAS includes documents such as: For the State's general assessments, evidence, such as a sections in the technical report for the assessments, that show: • A description of the process the State uses to ensure that the item types (e.g., multiple choice, constructed response, performance tasks, and technology-enhanced items) are tailored for assessing the academic content standards in terms of content; • A description of the process the State uses to ensure that items are tailored for assessing the academic content standards in terms of cognitive process (e.g., assessing complex demonstrations of knowledge and skills appropriate to the content, such as with item types that require synthesizing and evaluating information and analytical text-based writing or multiple steps and student explanations of their work); • Samples of item specifications that detail the content standards to be tested, item type, intended cognitive complexity, intended level of difficulty, accessibility tools and features, and response format; • Description or examples of instructions provided to item writers and reviewers; • Documentation that items are developed by individuals with content area expertise, experience as educators, and experience and expertise with students with disabilities, English learners, and other student populations in the State; • Documentation of procedures to review items for alignment to academic content standards, intended levels of cognitive complexity, intended levels of difficulty, construct-irrelevant variance, and consistency with item specifications, such as documentation of content and bias reviews by an external review committee;

	• Description of procedures to evaluate the quality of items and select items for operational use, including evidence of reviews of pilot and field test data; • As applicable, evidence that accessibility tools and features (e.g., embedded in test items or available as an accompaniment to the items) do not produce an inadvertent effect on the construct assessed; • Evidence that the items elicit the intended response processes, such as cognitive labs or interaction studies. AA-AAAS. For the State's AA-AAAS, in addition to the above: • If the State's AA-AAAS is a portfolio assessment, samples of item specifications that include documentation of the requirements for student work and samples of exemplars for illustrating levels of student performance; • Documentation of the process the State uses to ensure that the assessment items are accessible, cognitively challenging, and reflect professional judgment of the highest achievement standards possible for students with the most significant cognitive disabilities.  For the State's technology-based general assessments and AA-AAAS: • Documentation that procedures to evaluate and select items considered the deliverability of the items (e.g., usability studies). Note: This critical element is closely related to Critical Element 4.2 – Fairness and Accessibility.

Critical Element 2.3 – Test Administration

	Examples of Evidence

The State implements policies and procedures for standardized test administration, specifically the State:
• Has established and communicates to educators clear, thorough and consistent standardized procedures for the administration of its assessments, including administration with accommodations;
• Has established procedures to ensure that all individuals responsible for administering the State's general and alternate assessments receive training on the State's established procedures for the administration of its assessments;
	Evidence to support this critical element for the State's general assessments and AA-AAAS includes:
	
	• Regarding test administration:
	  ○ Test coordinator manuals, test administration manuals, accommodations manuals and/or other key documents that the State provides to districts, schools, and teachers that address standardized test administration and any accessibility tools and features available for the assessments;
	  ○ Instructions for the use of accommodations allowed by the State that address each accommodation. For example:
	  ▪ For accommodations such as bilingual dictionaries for English learners, instructions that indicate which types of bilingual dictionaries are and are not acceptable and how to acquire them for student use during the assessment;
	  ▪ For accommodations such as readers and scribes for students with disabilities, documentation of expectations for training and test security regarding test administration with readers and scribes;
	  ○ Evidence that the State provides key documents regarding test administration to district and school test coordinators and administrators, such as e-mails, websites, or listserv messages to inform relevant staff of the availability of documents for downloading or cover memos that accompany hard copies of the materials

• If the State administers technology-based assessments, the State has defined technology and other related requirements, included technology-based test administration in its standardized procedures for test administration, and established contingency plans to address possible technology challenges during test administration.	delivered to districts and schools; ○ Evidence of the State's process for documenting modifications or disruptions of standardized test administration procedures (e.g., unapproved non-standard accommodations, electric power failures or hardware failures during technology-based testing), such as sample of incidences documented during the most recent year of test administration in the State. • Regarding training for test administration: ○ Evidence regarding training, such as: ▪ Schedules for training sessions for different groups of individuals involved in test administration (e.g., district and school test coordinators, test administrators, school computer lab staff, accommodation providers); ▪ Training materials, such as agendas, slide presentations and school test coordinator manuals and test administrator manuals, provided to participants. For technology-based assessments, training materials that include resources such as practice tests and/or other supports to ensure that test coordinators, test administrators and others involved in test administration are prepared to administer the assessments; ▪ Documentation of the State's procedures to ensure that all test coordinators, test administrators, and other individuals involved in test administration receive training for each test administration, such as forms for sign-in sheets or screenshots of electronic forms for tracking attendance, assurance forms, or identification of individuals responsible for tracking attendance.  For the State's technology-based assessments: • Evidence that the State has clearly defined the technology (e.g., hardware, software, internet connectivity, and internet access) and other related requirements (e.g., computer lab configurations) necessary for schools to administer the assessments and has communicated these requirements to schools and districts; • District and school test coordinator manuals, test administrator manuals and/or other key documents that include specific instructions for administering technology-based assessments (e.g., regarding necessary advanced preparation, ensuring that test administrators and students are adequately familiar with the delivery devices and, as applicable, accessibility tools and features available for students); • Contingency plans or summaries of contingency plans that outline strategies for managing possible challenges or disruptions during test administration. **AA-AAAS.** For the State's AA-AAAS, in addition to the above: • If the assessments involve teacher-administered performance tasks or portfolios, key documents, such as test administration manuals, that the State provides to districts, schools and teachers that include clear, precise descriptions of activities, standard prompts, exemplars and scoring rubrics, as applicable; and standard procedures for the administration of the assessments that address features such as determining entry points, selection and use of manipulatives, prompts, scaffolding, and recognizing and recording responses; • Evidence that training for test administrators addresses key assessment features, such as teacher-administered performance tasks or portfolios; determining entry points; selection and use of manipulatives; prompts;

	scaffolding; recognizing and recording responses; and/or other features for which specific instructions may be needed to ensure standardized administration of the assessment.

Critical Element 2.4 – Monitoring Test Administration

	Examples of Evidence
The State adequately monitors the administration of its State assessments to ensure that standardized test administration procedures are implemented with fidelity across districts and schools.	Evidence to support this critical element for the State's general assessments and AA-AAAS includes documents such as: • Brief description of the State's approach to monitoring test administration (e.g., monitoring conducted by State staff, through regional centers, by districts with support from the State, or another approach); • Existing written documentation of the State's procedures for monitoring test administration across the State, including, for example, strategies for selection of districts and schools for monitoring, cycle for reaching schools and districts across the State, schedule for monitoring, monitors' roles, and the responsibilities of key personnel; • Summary of the results of the State's monitoring of the most recent year of test administration in the State.

Critical Element 2.5 – Test Security

	Examples of Evidence
The State has implemented and documented an appropriate set of policies and procedures to prevent test irregularities and ensure the integrity of test results through: • Prevention of any assessment irregularities, including maintaining the security of test materials, proper test preparation guidelines and administration procedures, incident-reporting procedures, consequences for confirmed violations of test security, and requirements for annual training at the district and school levels for all individuals involved in test administration; • Detection of test irregularities; • Remediation following any test security incidents involving any of	Collectively, evidence to support this critical element for the State's assessment system must demonstrate that the State has implemented and documented an appropriate approach to test security. Evidence to support this critical element for the State's assessment system may include: • State Test Security Handbook; • Summary results or reports of internal or independent monitoring, audit, or evaluation of the State's test security policies, procedures and practices, if any. Evidence of procedures for prevention of test irregularities includes documents such as: • Key documents, such as test coordinator manuals or test administration manuals for district and school staff, that include detailed security procedures for before, during and after test administration; • Documented procedures for tracking the chain of custody of secure materials and for maintaining the security of test materials at all stages, including distribution, storage, administration, and transfer of data; • Documented procedures for mitigating the likelihood of unauthorized communication, assistance, or recording of test materials (e.g., via technology such as smart phones); • Specific test security instructions for accommodations providers (e.g., readers, sign language interpreters, special education teachers and support staff if the assessment is administered individually), as applicable; • Documentation of established consequences for confirmed violations of test security, such as State law, State

the State's assessments; • Investigation of alleged or factual test irregularities.	regulations or State Board-approved policies; • Key documents such as policy memos, listserv messages, test coordinator manuals and test administration manuals that document that the State communicates its test security policies, including consequences for violation, to all individuals involved in test administration; • Newsletters, listserv messages, test coordinator manuals, test administrator manuals and/or other key documents from the State that clearly state that annual test security training is required at the district and school levels for all staff involved in test administration; • Evidence submitted under Critical Element 2.3 – Test Administration that shows: o The State's test administration training covers the relevant aspects of the State's test security policies; o Procedures for ensuring that all individuals involved in test administration receive annual test security training.  For the State's technology-based assessments, evidence of procedures for prevention of test irregularities includes: • Documented policies and procedures for districts and schools to address secure test administration challenges related to hardware, software, internet connectivity, and internet access. Evidence of procedures for detection of test irregularities includes documents such as: • Documented incident-reporting procedures, such as a template and instructions for reporting test administration irregularities and security incidents for district, school and other personnel involved in test administration; • Documentation of the information the State routinely collects and analyzes for test security purposes, such as description of post-administration data forensics analysis the State conducts (e.g., unusual score gains or losses, similarity analyses, erasure/answer change analyses, pattern analysis, person fit analyses, local outlier detection, unusual timing patterns); • Summary of test security incidents from most recent year of test administration (e.g., types of incidents and frequency) and examples of how they were addressed, or other documentation that demonstrates that the State identifies, tracks, and resolves test irregularities. Evidence of procedures for remediation of test irregularities includes documents such as: • Contingency plan that demonstrates that the State has a plan for how to respond to test security incidents and that addresses: o Different types of possible test security incidents (e.g., human, physical, electronic, or internet-related), including those that require immediate action (e.g., items exposed on-line during the testing window); o Policies and procedures the State would use to address different types of test security incidents (e.g., continue vs. stop testing, retesting, replacing existing forms or items, excluding items from scoring, invalidating results); o Communication strategies for communicating with districts, schools and others, as appropriate, for addressing active events.

	Evidence of procedures for investigation of alleged or factual test irregularities includes documents such as: • State's policies and procedures for responding to and investigating, where appropriate, alleged or actual security lapses and test irregularities that: o Include securing evidence in cases where an investigation may be pursued; o Include the State's decision rules for investigating potential test irregularities; o Provide standard procedures and strategies for conducting investigations, including guidelines to districts, if applicable; o Include policies and procedures to protect the privacy and professional reputation of all parties involved in an investigation. Note: Evidence should be redacted to protect personally identifiable information, as appropriate.

Critical Element 2.6 – Systems for Protecting Data Integrity and Privacy

	Examples of Evidence
The State has policies and procedures in place to protect the integrity and confidentiality of its test materials, test-related data, and personally identifiable information, specifically: • To protect the integrity of its test materials and related data in test development, administration, and storage and use of results; • To secure student-level assessment data and protect student privacy and confidentiality, including guidelines for districts and schools; • To protect personally identifiable information about any individual student in reporting, including defining the minimum number of students necessary to allow reporting of scores for all students and student groups.	Evidence to support this critical element for the State's general assessments and AA-AAAS includes documents such as: • Evidence of policies and procedures to protect the integrity and confidentiality of test materials and test-related data, such as: o State security plan, or excerpts from the State's assessment contracts or other materials that show expectations, rules and procedures for reducing security threats and risks and protecting test materials and related data during item development, test construction, materials production, distribution, test administration, and scoring; o Description of security features for storage of test materials and related data (i.e., items, tests, student responses, and results); o Rules and procedures for secure transfer of student-level assessment data in and out of the State's data management and reporting systems; between authorized users (e.g., State, district and school personnel, and vendors); and at the local level (e.g., requirements for use of secure sites for accessing data, directions regarding the transfer of student data); o Policies and procedures for allowing only secure, authorized access to the State's student-level data files for the State, districts, schools, and others, as applicable (e.g., assessment consortia, vendors); o Training requirements and materials for State staff, contractors and vendors, and others related to data integrity and appropriate handling of personally identifiable information; o Policies and procedures to ensure that aggregate or de-identified data intended for public release do not inadvertently disclose any personally identifiable information; o Documentation that the above policies and procedures, as applicable, are clearly communicated to all relevant personnel (e.g., State staff, assessment, districts, and schools, and others, as applicable (e.g., assessment consortia, vendors)); o Rules and procedures for ensuring that data released by third parties (e.g., agency partners, vendors,

	external researchers) are reviewed for adherence to State Statistical Disclosure Limitation (SDL) standards and do not reveal personally identifiable information.
	
	• Evidence of policies and procedures to protect personally identifiable information about any individual student in reporting, such as:
	  ○ State operations manual or other documentation that clearly states the State's SDL rules for determining whether data are reported for a group of students or a student group, including:
	  ▪ Defining the minimum number of students necessary to allow reporting of scores for a student group;
	  ▪ Rules for applying complementary suppression (or other SDL methods) when one or more student groups are not reported because they fall below the minimum reporting size;
	  ▪ Rules for not reporting results, regardless of the size of the student group, when reporting would reveal personally identifiable information (e.g., procedures for reporting "<10%" for proficient and above when no student scored at those levels);
	  ▪ Other rules to ensure that aggregate or de-identified data do not inadvertently disclose any personally identifiable information;
	  ○ State operations manual or other document that describes how the State's rules for protecting personally identifiable information are implemented.

SECTION 3: TECHNICAL QUALITY – VALIDITY

Critical Element 3.1 – Overall Validity, Including Validity Based on Content

	Examples of Evidence
The State has documented adequate overall validity evidence for its assessments, and the State's validity evidence includes evidence that the State's assessments measure the knowledge and skills specified in the State's academic content standards, including: • Documentation of adequate alignment between the State's assessments and the academic content standards the assessments are designed to measure in terms of content (i.e., knowledge and process), the full range of the State's academic content standards, balance of content, and cognitive complexity; • If the State administers alternate assessments based on alternate academic achievement standards, the assessments show adequate linkage to the State's academic content standards in terms of content match (i.e., no unrelated content) and the breadth of content and cognitive complexity determined in test design to be appropriate for students with the most significant cognitive disabilities.	Collectively, across the State's assessments, evidence to support critical elements 3.1 through 3.4 for the State's general assessments and AA-AAAS must document overall validity evidence generally consistent with expectations of current professional standards. Evidence to document adequate overall validity evidence for the State's general assessments and AA-AAAS includes documents such as: • A chapter on validity in the technical report for the State's assessments that states the purposes of the assessments and intended interpretations and uses of results and shows validity evidence for the assessments that is generally consistent with expectations of current professional standards; • Other validity evidence, in addition to that outlined in critical elements 3.1 through 3.4, that is necessary to document adequate validity evidence for the assessments. Evidence to document adequate validity evidence based on content for the State's general assessments includes: • Validity evidence based on the assessment content that shows levels of validity generally consistent with expectations of current professional standards, such as: o Test blueprints, as submitted under Critical Element 2.1—Test Design and Development; o A full form of the assessment in one grade for the general assessment in reading/language arts and mathematics (e.g., one form of the grade 5 mathematics assessment and one form of the grade 8 reading/language arts assessment);[6] o Logical or empirical analyses that show that the test content adequately represents the full range of the State's academic content standards; o Report of expert judgment of the relationship between components of the assessment and the State's academic content standards; o Reports of analyses to demonstrate that the State's assessment content is appropriately related to the specific inferences made from test scores about student proficiency in the State's academic content standards for all student groups; • Evidence of alignment, including: o Report of results of an independent alignment study that is technically sound (i.e., method and process, appropriate units of analysis, clear criteria) and documents adequate alignment, specifically that: ▪ Each assessment is aligned to its test blueprint, and each blueprint is aligned to the full range of State's

[6] The Department recognizes the need for a State to maintain the security of its test forms; a State that elects to submit a test form(s) as part of its assessment peer review submission should contact the Department so that arrangements can be made to ensure that the security of the materials is maintained. Such materials will be reviewed by the assessment peer reviewers in accordance with the State's test security requirements and agreements.

	academic content standards; or ▪ Each assessment is aligned to the full range of the State's academic content standards, and the procedures the State follows to ensure such alignment during test development; o Description of a systematic process and timeline the State will implement to address any gaps or weaknesses identified in the alignment studies.  For the State's computer-adaptive general assessments: • Empirical evidence that the size of the item pool and the characteristics (non-statistical (e.g., content) and statistical) of items it contains are appropriate for the test design and adequately reflect the blueprint in terms of: o Full range of the State's grade-level academic content standards; o Balance of content; o Cognitive complexity for each standard tested; o Range of item difficulty levels for each standard tested; o Structure of the assessment (e.g., number of items and proportion of item and response types specified by the blueprints); o Item pool size and composition sufficient to avoid over-exposure of items; • Results of an alignment study confirming that the test forms generated for individual students are aligned to the State's academic content standards in terms of: o Full range of the State's grade-level academic content standards; o Balance of content; o Cognitive complexity for each standard tested; o Range of item difficulty levels for each standard tested; o Structure of the assessment (i.e., features specified in Critical Element 2.1 – Test Design and Development, such as number of items and proportion of item and response types specified by the blueprints); • Empirical analyses that show: o The actual test forms produce an adequately precise estimate of student achievement; o Students are appropriately routed to the next item or stage based on their responses to the previous item or stage; o Response data adequately fit the psychometric model selected by the State. **AA-AAAS.** For the State's AA-AAAS, evidence to document adequate validity evidence based on content includes: • Validity evidence that shows levels of validity generally considered adequate by professional judgment regarding such assessments, such as: o Test blueprints and other evidence submitted under Critical Element 2.1 – Test Design and Development; o Evidence documenting adequate linkage between the assessments and the academic content they are intended to measure; o Other documentation that shows the State's assessments measure only the knowledge and skills specified in the State's academic content standards (or extended academic content standards, as applicable) for the tested grade (i.e., not unrelated content);

	• Evidence of alignment, such as:
	  ○ Report of results of an independent alignment study that is technically sound and document adequate linkage between each of the State's assessments and the academic content the assessments are designed to measure;
	  ○ If the State developed extended academic content standards for students with the most significant cognitive disabilities and used these to develop its AA-AAAS, the alignment study should document the linkage between the State's academic content standards and extended academic content standards as well as adequate linkage between the extended academic content standards and the assessments;
	• For an adaptive AA-AAAS:
	  ○ Summary of an analysis to confirm that the item pool adequately represents the test blueprints, such as a crosswalk of the item pool and the test blueprints;
	  ○ Results of an alignment study that confirm that the test design, as implemented, produces assessments with adequate linkage to the academic content standards the assessments are designed to measure.

Critical Element 3.2 – Validity Based on Cognitive Processes

	Examples of Evidence

The State has documented adequate validity evidence that its assessments tap the intended cognitive processes appropriate for each grade level as represented in the State's academic content standards.	Evidence to support this critical element for the State's general assessments includes:
	• Validity evidence based on cognitive processes that shows levels of validity generally consistent with expectations of current professional standards, such as:
	  ○ Results of cognitive labs exploring student performance on items that show the items require complex demonstrations or applications of knowledge and skills;
	  ○ Reports of expert judgment of items that show the items require complex demonstrations or applications of knowledge and skills;
	  ○ Empirical evidence that shows the relationships of items intended to require complex demonstrations or applications of knowledge and skills to other measures that require similar levels of cognitive complexity in the content area (e.g., teacher ratings of student performance, student performance on performance tasks or external assessments of the same knowledge and skills).
	
	AA-AAAS. For the State's AA-AAAS, evidence to support this critical element includes:
	• Validity evidence that shows levels of validity generally considered adequate by professional judgment regarding such assessments, such as:
	  ○ Results of cognitive labs exploring student performance on items that show the items require demonstrations or applications of knowledge and skills;
	  ○ Reports of expert judgment of items that show the items require demonstrations or applications of knowledge and skills;
	  ○ Empirical evidence that shows the relationships of items intended to require demonstrations or applications of knowledge and skills to other measures that require similar levels of cognitive complexity in the content

	area (e.g., teacher ratings of student performance, student performance on performance tasks or external assessments of the same knowledge and skills).

Critical Element 3.3 – Validity Based on Internal Structure

	Examples of Evidence
The State has documented adequate validity evidence that the scoring and reporting structures of its assessments are consistent with the sub-domain structures of the State's academic content standards on which the intended interpretations and uses of results are based.	Evidence to support this critical element for the State's general assessments includes: • Validity evidence based on the internal structure of the assessments that shows levels of validity generally consistent with expectations of current professional standards, such as: o Reports of analyses of the internal structure of the assessments (e.g., tables of item correlations) that show the extent to which the interrelationships among subscores are consistent with the State's academic content standards for relevant student groups; o Reports of analyses that show the dimensionality of the assessment is consistent with the structure of the State's academic content standards and the intended interpretations of results; o Evidence that ancillary constructs needed for success on the assessments do not provide inappropriate barriers for measuring the achievement of all students, such as evidence from cognitive labs or documentation of item development procedures; o Reports of differential item functioning (DIF) analyses that show whether particular items (e.g., essays, performance tasks, or items requiring specific knowledge or skills) function differently for relevant student groups. **AA-AAAS.** For the State's AA-AAAS, evidence to support this critical element includes: • Validity evidence that shows levels of validity generally considered adequate by professional judgment regarding such assessments, such as: o Validity evidence based on the internal structure of the assessments, such as analysis of response patterns for administered items (e.g., student responses indicating no attempts at answering questions or suggesting guessing).

Critical Element 3.4 – Validity Based on Relations to Other Variables

	Examples of Evidence

The State has documented adequate validity evidence that the State's assessment scores are related as expected with other variables.	Evidence to support this critical element for the State's general assessments includes:
	• Validity evidence that shows the State's assessment scores are related as expected with criterion and other variables for all student groups, such as:
	  ○ Reports of analyses that demonstrate positive correlations between State assessment results and external measures that assess similar constructs, such as NAEP, TIMSS, assessments of the same content area administered by some or all districts in the State, and college-readiness assessments;
	  ○ Reports of analyses that demonstrate convergent relationships between State assessment results and measures other than test scores, such as performance criteria, including college- and career-readiness (e.g., college-enrollment rates; success in related entry-level, college credit-bearing courses; post-secondary employment in jobs that pay living wages);
	  ○ Reports of analyses that demonstrate positive correlations between State assessment results and other variables, such as academic characteristic of test takers (e.g., average weekly hours spent on homework, number of advanced courses taken);
	  ○ Reports of analyses that show stronger positive relationships with measures of the same construct than with measures of different constructs;
	  ○ Reports of analyses that show assessment scores at tested grades are positively correlated with teacher judgments of student readiness at entry in the next grade level.
	
	AA-AAAS. For the State's AA-AAAS, evidence to support this critical element includes:
	• Validity evidence that shows levels of validity generally considered adequate by professional judgment regarding such assessments, such as:
	  ○ Validity evidence based on relationships with other variables, such as analyses that demonstrate positive correlations between assessment results and other variables, for example:
	  ▪ Correlations between assessment results and variables related to test takers (e.g., instructional time on content based on grade-level content standards);
	  ▪ Correlations between proficiency on the high-school assessments and performance in post-secondary education, vocational training or employment.

SECTION 4: TECHNICAL QUALITY – OTHER

Critical Element 4.1 – Reliability

	Examples of Evidence

The State has documented adequate reliability evidence for its assessments for the following measures of reliability for the State's student population overall and each student group and, if the State's assessments are implemented in multiple States, for the assessment overall and each student group, including:
• Test reliability of the State's assessments estimated for its student population;
• Overall and conditional standard error of measurement of the State's assessments;
• Consistency and accuracy of estimates in categorical classification decisions for the cut scores and achievement levels based on the assessment results;
• For computer-adaptive tests, evidence that the assessments produce test forms with adequately precise estimates of a student's achievement.
	Collectively, evidence for the State's general assessments and AA-AAAS must document adequate reliability evidence generally consistent with expectations of current professional standards.
	
	Evidence to support this critical element for the State's general assessments includes documentation such as:
	• A chapter on reliability in the technical report for the State's assessments that shows reliability evidence;
	• For the State's general assessments, documentation of reliability evidence generally consistent with expectations of current professional standards, including:
	  ○ Results of analyses for alternate-form or, test-retest internal consistency reliability statistics, as appropriate, for each assessment;
	  ○ Report of standard errors of measurement and conditional standard errors of measurement, for example, in terms of one or more coefficients or IRT-based test information functions at each cut score specified in the State's academic achievement standards;
	  ○ Results of estimates of decision consistency and accuracy for the categorical decisions (e.g., classification of proficiency levels) based on the results of the assessments.
	
	For the State's computer-adaptive assessments, evidence that estimates of student achievement are adequately precise includes documentation such as:
	• Summary of empirical analyses showing that the estimates of student achievement are adequately precise for the intended interpretations and uses of the student's assessment score;
	• Summary of analyses that demonstrates that the test forms are adequately precise across all levels of ability in the student population overall and for each student group (e.g., analyses of the test information functions and conditional standard errors of measurement).
	
	AA-AAAS. For the State's AA-AAAS, evidence to support this critical element includes:
	• Reliability evidence that shows levels of reliability generally considered adequate by professional judgment regarding such assessments includes documentation such as:
	  ○ Internal consistency coefficients that show that item scores are related to a student's overall score;
	  ○ Correlations of item responses to student proficiency level classifications;
	  ○ Generalizability evidence such as evidence of fidelity of administration;
	  ○ As appropriate and feasible given the size of the tested population, other reliability evidence as outlined above.

Critical Element 4.2 – Fairness and Accessibility

	Examples of Evidence

The State has taken reasonable and appropriate steps to ensure that its assessments are accessible to all students and fair across student groups in the design, development and analysis of its assessments.	Evidence to support this critical element for the State's general assessments and AA-AAAS includes:
	
	For the State's general assessments:
	• Documentation of steps the State has taken in the design and development of its assessments, such as:
	  ○ Documentation describing approaches used in the design and development of the State's assessments (e.g., principles of universal design, language simplification, accessibility tools and features embedded in test items or available as an accompaniment to the items);
	  ○ Documentation of the approaches used for developing items;
	  ○ Documentation of procedures used for maximizing accessibility of items during the development process, such as guidelines for accessibility and accessibility tools and features included in item specifications;
	  ○ Description or examples of instructions provided to item writers and reviewers that address writing accessible items, available accessibility tools and features, and reviewing items for accessibility;
	  ○ Documentation of procedures for developing and reviewing items in alternative formats or substitute items and for ensuring these items conform with item specifications;
	  ○ Documentation of routine bias and sensitivity training for item writers and reviewers;
	  ○ Documentation that experts in the assessment of students with disabilities, English learners and individuals familiar with the needs of other student populations in the State were involved in item development and review;
	  ○ Descriptions of the processes used to write, review, and evaluate items for bias and sensitivity;
	  ○ Description of processes to evaluate items for bias during pilot and field testing;
	  ○ Evidence submitted under Critical Elements 2.1 – Test Design and Development and Critical Element 2.2 – Item Development;
	• Documentation of steps the State has taken in the analysis of its assessments, such as results of empirical analyses (e.g., DIF and differential test functioning (DTF) analyses) that identify possible bias or inconsistent interpretations of results across student groups.
	
	AA-AAAS. For the State's AA-AAAS:
	• Documentation of steps the State has taken in the design and development of its assessments, as listed above;
	• Documentation of steps the State has taken in the analysis of its assessments, for example:
	  ○ Results of bias reviews or, when feasible given the size of the tested student population, empirical analyses (e.g., DIF analyses and DTF analyses by disability category);
	  ○ Frequency distributions of the tested population by disability category;
	  ○ As appropriate, applicable and feasible given the size of the tested population, other evidence as outlined above.
	
	Note: This critical element is closely related to Critical Element 2.2 – Item Development.

Critical Element 4.3 – Full Performance Continuum

	Examples of Evidence
The State has ensured that each assessment provides an adequately precise estimate of student performance across the full performance continuum, including for high- and low-achieving students.	Evidence to support this critical element for the State's general assessments and AA-AAAS includes documents such as: For the State's general assessments: • Description of the distribution of cognitive complexity and item difficulty indices that demonstrate the items included in each assessment adequately cover the full performance continuum; • Analysis of test information functions (TIF) and ability estimates for students at different performance levels across the full performance continuum or a pool information function across the full performance continuum; • Table of conditional standard errors of measurement at various points along the score range. AA-AAAS. For the State's AA-AAAS: • A cumulative frequency distribution or histogram of student scores for each grade and subject on the most recent administration of the State's assessment; • For students at the lowest end of the performance continuum (e.g., pre-symbolic language users or students with no consistent communicative competencies), evidence that the assessments provide appropriate performance information (e.g., communicative competence); • As appropriate, applicable and feasible given the size of the tested population, other evidence as outlined above.

Critical Element 4.4 – Scoring

	Examples of Evidence

The State has established and documented standardized scoring procedures and protocols for its assessments that are designed to produce reliable results, facilitate valid score interpretations, and report assessment results in terms of the State's academic achievement standards.	Evidence to support this critical element for the State's general assessments and AA-AAAS includes:
	• A chapter on scoring in a technical report for the assessments or other documentation that describes scoring procedures, including:
	  ○ Procedures for constructing scales used for reporting scores and the rationale for these procedures;
	  ○ Scale, measurement error, and descriptions of test scores;
	• For scoring involving human judgment:
	  ○ Evidence that the scoring of constructed-response items and performance tasks includes adequate procedures and criteria for ensuring and documenting inter-rater reliability (e.g., clear scoring rubrics, adequate training for and qualifying of raters, evaluation of inter-rater reliability, and documentation of quality control procedures);
	  ○ Results of inter-rater reliability of scores on constructed-response items and performance tasks;
	• For machine scoring of constructed-response items:
	  ○ Evidence that the scoring algorithm and procedures are appropriate, such as descriptions of development and calibration, validation procedures, monitoring, and quality control procedures;

	○ Evidence that machine scoring produces scores are comparable to those produced by human scorers, such as rater agreement rates for human- and machine-scored samples of responses (e.g., by student characteristics such as varying achievement levels and student groups), systematic audits and rescores; • Documentation that the system produces student results in terms of the State's academic achievement standards; • Documentation that the State has rules for invalidating test results when necessary (e.g., non-attempt, cheating, unauthorized accommodation or modification) and appropriate procedures for implementing these rules (e.g., operations manual for the State's assessment and accountability systems, test coordinators manuals and test administrator manuals, or technical reports for the assessments). **AA-AAAS.** For the State's AA-AAAS, in addition to the above: • If the assessments are portfolio assessments, evidence of procedures to ensure that only student work including content linked to the State's grade-level academic content standards is scored; • If the alternate assessments involve any scoring of performance tasks by test administrators (e.g., teachers): ○ Evidence of adequate training for all test administrators (may include evidence submitted under Critical Element 2.3 – Test Administration); ○ Procedures the State uses for each test administration to ensure the reliability of scoring; ○ Documentation of the inter-rater reliability of scoring by test administrators.

Critical Element 4.5 – Multiple Assessment Forms

	Examples of Evidence
If the State administers multiple forms within a content area and grade level, within or across school years, the State ensures that all forms adequately represent the State's academic content standards and yield consistent score interpretations such that the forms are comparable within and across school years.	Evidence to support this critical element for the State's assessments system includes documents such as: • Documentation of technically sound equating procedures and results within an academic year, such as a section of a technical report for the assessments that provides detailed technical information on the method used to establish linkages and on the accuracy of equating functions; • As applicable, documentation of year-to-year equating procedures and results, such as a section of a technical report for the assessments that provides detailed technical information on the method used to establish linkages and on the accuracy of equating functions.

Critical Element 4.6 – Multiple Versions of an Assessment

	Examples of Evidence
If the State administers assessments in multiple versions within a content area, grade level, or school year, the State:	Evidence to support this critical element for the State's general and alternate assessments includes: For the State's general assessments:

• Followed a design and development process to support comparable interpretations of results for students tested across the versions of the assessments;
• Documented adequate evidence of comparability of the meaning and interpretations of the assessment results.
	• Documentation that the State followed a design and development process to support comparable interpretations of results across different versions of the assessments (e.g., technology-based and paper-based assessments, assessments in English and native language(s), general and alternate assessments based on grade-level academic achievement standards);
	  ○ For a native language assessment, this may include a description of the State's procedures for translation or trans-adaptation of the assessment or a report of analysis of results of back-translation of a translated test;
	  ○ For technology-based and paper-based assessments, this may include demonstration that the provision of paper-based substitutes for technology-enabled items elicits comparable response processes and produces an adequately aligned assessment;
	• Report of results of a comparability study of different versions of the assessments that is technically sound and documents evidence of comparability generally consistent with expectations of current professional standards.
	
	  If the State administers technology-based assessments that are delivered by different types of devices (e.g., desktop computers, laptops, tablets), evidence includes:
	• Documentation that test-administration hardware and software (e.g., screen resolution, interface, input devices) are standardized across unaccommodated administrations; or
	• Either:
	  ○ Reports of research (quantitative or qualitative) that show that variations resulting from different types of delivery devices do not alter the interpretations of results; or
	  ○ A comparability study, as described above.
	
	AA-AAAS. For the State's AA-AAAS:
	• Documentation that the State followed design, development and test administration procedures to ensure comparable results across different versions of the assessments, such as a description of the processes in the technical report for the assessments or a separate report.

Critical Element 4.7 – Technical Analysis and Ongoing Maintenance

	Examples of Evidence

The State has a system for monitoring and maintaining, and improving as needed, the quality of its assessment system, including clear and technically sound criteria for the analyses of all of the assessments in its assessment system (i.e., general assessments and alternate assessments).	Evidence to support this critical element for the State's assessments system includes:
	• Documentation that the State has established and implemented clear and technically sound criteria for analyses of its assessment system, such as:
	  ○ Sections from the State's assessment contract that specify the State's expectations for analyses to provide evidence of validity, reliability, and fairness; for independent studies of alignment and comparability, as appropriate; and for requirements for technical reports for the assessments and the content of such reports applicable to each administration of the assessment;
	  ○ The most recent technical reports for the State's assessments that present technical analyses of the State's

	assessments; o Documentation of the alignment of the State's assessments to the State's academic content standards (e.g., evidence submitted under Critical Element 3.1 – Overall Validity, Including Validity Based on Content; o Presentations of assessments results (e.g., to the State's TAC); • Documentation of the State's system for monitoring and improving, as needed, the on-going quality of its assessment system, such as: o Evidence that the State has established and implemented clear criteria for the analysis of its assessment system (see above); o Documentation of regular internal and external technical review of components of the State's assessment system, such as State Board of Education minutes, minutes from TAC meetings, and documentation of roles and responsibilities of TAC members; o Outline of a deliberate cycle for reviewing and updating the State's academic content standards and assessments (e.g., provides for logical transitions such that the assessments are aligned to the standards on which instruction is based in the relevant school year).

SECTION 5: INCLUSION OF ALL STUDENTS

Critical Element 5.1 – Procedures for Including Students with Disabilities

	Examples of Evidence
The State has in place procedures to ensure the inclusion of all public elementary and secondary school students with disabilities in the State's assessment system, including, at a minimum, guidance for IEP Teams to inform decisions about student assessments that: • Provides clear explanations of the differences between assessments based on grade-level academic achievement standards and assessments based on alternate academic achievement standards, including any effects of State and local policies on a student's education resulting from taking an alternate assessment based on alternate academic achievement standards; • States that decisions about how to assess students with disabilities must be made by a student's IEP Team based on each student's individual needs; • Provides guidelines for determining whether to assess a student on the general assessment without accommodation(s), the general assessment with accommodation(s), or an alternate assessment; • Provides information on accessibility tools and features available to students in general and assessment accommodations available for students with disabilities;	Evidence to support this critical element for the State's assessment system includes: • Documentation that the State has in place procedures to ensure the inclusion of all students with disabilities, such as: o Guidance for IEP Teams and IEP templates for students in tested grades; o Training materials for IEP Teams; o Accommodations manuals or other key documents that provide information on accommodations for students with disabilities; o Test administration manuals or other key documents that provide information on available accessibility tools and features; • Documentation that the implementation of the State's alternate academic achievement standards promotes student access to the general curriculum, such as: o State policies that require that instruction for students with the most significant cognitive disabilities be linked to the State's grade-level academic content standards; o State policies that require standards-based IEPs linked to the State's grade-level academic content standards for students with the most significant cognitive disabilities; o Reports of State monitoring of IEPs that document the implementation of IEPs linked to the State's grade-level academic content standards for students with the most significant cognitive disabilities. Note: Key topics related to the assessment of students with disabilities are also addressed in Critical Element 4.2 -- Fairness and Accessibility and in critical elements addressing the AA-AAAS throughout.

• Provides guidance regarding selection of appropriate accommodations for students with disabilities; • Includes instructions that students eligible to be assessed based on alternate academic achievement standards may be from any of the disability categories listed in the IDEA; • Ensures that parents of students with the most significant cognitive disabilities are informed that their student's achievement will be based on alternate academic achievement standards and of any possible consequences of taking the alternate assessments resulting from district or State policy (e.g., ineligibility for a regular high school diploma if the student does not demonstrate proficiency in the content area on the State's general assessments); • The State has procedures in place to ensure that its implementation of alternate academic achievement standards for students with the most significant cognitive disabilities promotes student access to the general curriculum.	

Critical Element 5.2 – Procedures for Including English Learners

	Examples of Evidence
The State has in place procedures to ensure the inclusion of all English learners in public elementary and secondary schools in the State's assessment system and clearly	Evidence to support this critical element for the State's assessment system includes: • Documentation of procedures for determining student eligibility for accommodations and guidance on selection of appropriate accommodations for English learners; • Accommodations manuals or other key documents that provide information on accommodations for English learners;

communicates this information to districts, schools, teachers, and parents, including, at a minimum: • Procedures for determining whether an English learner should be assessed with accommodation(s); • Information on accessibility tools and features available to all students and assessment accommodations available for English learners; • Guidance regarding selection of appropriate accommodations for English learners.	• Test administration manuals or other key documents that provide information on available accessibility tools and features; • Guidance in key documents that indicates all accommodation decisions must be based on individual student needs and provides suggestions regarding what types of accommodations may be most appropriate for students with various levels of proficiency in their first language and English. Note: Key topics related to the assessment of English learners are also addressed in Critical Element 4.2 – Fairness and Accessibility.

Critical Element 5.3 – Accommodations

	Examples of Evidence
The State makes available appropriate accommodations and ensures that its assessments are accessible to students with disabilities and English learners. Specifically, the State: • Ensures that appropriate accommodations are available for students with disabilities under IDEA and students covered by Section 504; • Ensures that appropriate accommodations are available for English learners; • Has determined that the accommodations it provides (i) are appropriate and effective for meeting the individual student's need(s) to participate in the assessments, (ii) do not alter the construct being assessed, and (iii) allow meaningful interpretations of results and comparison of scores for students who need and receive	Evidence to support this critical element for both the State's general and AA-AAAS includes: • Lists of accommodations available for students with disabilities under IDEA, students covered by Section 504 and English learners that are appropriate and effective for addressing barrier(s) faced by individual students (i.e., disability and/or language barriers) and appropriate for the assessment mode (e.g., paper-based vs. technology-based), such as lists of types of available accommodations in an accommodations manual, test coordinators manual or test administrators manual; • Documentation that scores for students based on assessments administered with allowable accommodations (and accessibility tools and features, as applicable) allow for valid inferences, such as: o Description of the reasonable and appropriate basis for the set of accommodations offered on the assessments, such as a literature review, empirical research, recommendations by advocacy and professional organizations, and/or consultations with the State's TAC, as documented in a section on test design and development in the technical report for the assessments; o For accommodations not commonly used in large-scale State assessments, not commonly used in the manner adopted for the State's assessment system, or newly developed accommodations, reports of studies, data analyses, or other evidence that indicate that scores based on accommodated and non-accommodated administrations can be meaningfully compared; o A summary of the frequency of use of each accommodation on the State's assessments by student characteristics (e.g., students with disabilities, English learners); • Evidence that the State has a process to review and approve requests for assessment accommodations beyond those routinely allowed, such as documentation of the State's process as communicated to district and school test coordinators and test administrators.

accommodations and students who do not need and do not receive accommodations; • Has a process to individually review and allow exceptional requests for a small number of students who require accommodations beyond those routinely allowed.	

Critical Element 5.4 – Monitoring Test Administration for Special Populations

	Examples of Evidence
The State monitors test administration in its districts and schools to ensure that appropriate assessments, with or without appropriate accommodations, are selected for students with disabilities under IDEA, students covered by Section 504, and English learners so that they are appropriately included in assessments and receive accommodations that are: • Consistent with the State's policies for accommodations; • Appropriate for addressing a student's disability or language needs for each assessment administered; • Consistent with accommodations provided to the students during instruction and/or practice; • Consistent with the assessment accommodations identified by a student's IEP Team or 504 team for students with disabilities, or another process for an English learner; • Administered with fidelity to test administration procedures.	Evidence to support this critical element for the State's assessment system includes documents such as: • Description of procedures the State uses to monitor that accommodations selected for students with disabilities, students covered by Section 504, and English learners are appropriate; • Description of procedures the State uses to monitor that students with disabilities are placed by IEP Teams in the appropriate assessment; • The State's written procedures for monitoring the use of accommodations during test administration, such as guidance provided to districts; instructions and protocols for State, district and school staff; and schedules for monitoring; • Summary of results of monitoring for the most recent year of test administration in the State.

SECTION 6: ACADEMIC ACHIEVEMENT STANDARDS AND REPORTING

Critical Element 6.1 – State Adoption of Academic Achievement Standards for All Students

	Examples of Evidence
The State formally adopted challenging academic achievement standards in reading/language arts, mathematics and in science for all students, specifically: • The State formally adopted academic achievement standards in the required tested grades and, at its option, also alternate academic achievement standards for students with the most significant cognitive disabilities; • The State applies its grade-level academic achievement standards to all public elementary and secondary school students enrolled in the grade to which they apply, with the exception of students with the most significant cognitive disabilities to whom alternate academic achievement standards may apply; • The State's academic achievement standards and, as applicable, alternate academic achievement standards, include: (a) At least three levels of achievement, with two for high achievement and a third for lower achievement; (b) descriptions of the competencies associated with each achievement level; and (c) achievement scores that differentiate among the achievement levels.	Evidence to support this critical element for the State's assessment system includes: • Evidence of adoption of the State's academic achievement standards and, as applicable, alternate academic achievement standards, in the required tested grades and subjects (i.e., in reading/language arts and mathematics for each of grades 3-8 and high school and in science for each of three grade spans (3-5, 6-9, and 10-12)), such as State Board of Education minutes, memo announcing formal approval from the Chief State School Officer to districts, legislation, regulations, or other binding approval of academic achievement standards and, as applicable, alternate academic achievement standards; • State statutes, regulations, policy memos, State Board of Education minutes, memo from the Chief State School Officer to districts or other key documents that clearly state that the State's academic achievement standards apply to all public elementary and secondary school students in the State (with the exception of students with the most significant cognitive disabilities to whom alternate academic achievement standards may apply); • Evidence regarding the academic achievement standards and, as applicable, alternate academic achievement standards, regarding: (a) at least three levels of achievement, including two levels of high achievement (e.g., proficient and advanced) and a third of lower achievement (e.g., basic); (b) descriptions of the competencies associated with each achievement level; and (c) achievement scores (i.e., "cut scores") that differentiate among the achievement levels.

Critical Element 6.2 – Achievement Standards Setting

	Examples of Evidence
The State used a technically sound method and process that involved panelists with appropriate experience and expertise for setting its academic achievement standards and alternate academic achievement standards to ensure they are valid and reliable.	Evidence to support this critical element for the State's general assessments and AA-AAAS includes: • The State's standards-setting report, including: o A description of the standards-setting method and process used by the State; o The rationale for the method selected; o Documentation that the method used for setting cut scores allowed panelists to apply their knowledge and experience in a reasonable manner and supported the establishment of reasonable and defensible cut scores; o Documentation of the process used for setting cut scores and developing performance-level descriptors aligned to the State's academic content standards; o A description of the process for selecting panelists; o Documentation that the standards-setting panels consisted of panelists with appropriate experience and expertise, including: ▪ Content experts with experience teaching the State's academic content standards in the tested grades; ▪ Individuals with experience and expertise teaching students with disabilities, English learners and other student populations in the State; ▪ As appropriate, individuals from institutions of higher education (IHE) and individuals knowledgeable about career-readiness; ▪ A description, by relevant characteristics, of the panelists (overall and by individual panels) who participated in achievement standards setting; o If available, a summary of statistical descriptions and analyses that provides evidence of the reliability of the cut scores and the validity of recommended interpretations. **AA-AAAS.** For the State's AA-AAAS, in addition to the above: • Documentation that the panels for setting alternate academic achievement standards included individuals knowledgeable about the State's academic content standards and special educators knowledgeable about students with the most significant cognitive disabilities.

Critical Element 6.3 – Challenging and Aligned Academic Achievement Standards

	Examples of Evidence
The State's academic achievement standards are challenging and aligned with the State's academic content standards such that a high school student who scores at the proficient or above level has mastered what students are expected to know and be able to do by the time	Evidence to support this critical element for the State's general assessments and AA-AAAS includes: For the State's general assessments: • Documentation that the State's academic achievement standards are aligned with the State's academic content standards, such as: o A description of the process used to develop the State's academic achievement standards that shows that: ▪ The State's grade-level academic content standards were used as a main reference in writing

they graduate from high school in order to succeed in college and the workforce. If the State has defined alternate academic achievement standards for students with the most significant cognitive disabilities, the alternate academic achievement standards are linked to the State's grade-level academic content standards or extended academic content standards, show linkage to different content across grades, and reflect professional judgment of the highest achievement standards possible for students with the most significant cognitive disabilities.	performance level descriptors; ▪ The process of setting cut scores used, as a main reference, performance level descriptors that reflect the State's grade-level academic content standards; ▪ The State's cut scores were set and performance level descriptors written to reflect the full range of the State's academic content standards for each grade; o A description of steps taken to vertically articulate the performance level descriptors across grades; o Evaluation by standard-setting panelists or external expert reviewers that the State's academic achievement standards are aligned to the grade-level academic content standards and include subject-specific performance level descriptors that meaningfully differentiate across performance levels within grades and are vertically articulated across grades; • Documentation that the State's academic achievement standards are challenging, such as: o Reports of the results of benchmarking the State's academic achievement standards against NAEP, international assessments or other related and appropriate measures; o Policies of the State network of institutions of higher education (IHEs) that exempt from remedial courses and place into credit-bearing college courses any student who scores at the proficient level or above on the State's high school assessments. **AA-AAAS.** For the State's AA-AAAS: • Documentation that the State's alternate academic achievement standards are linked to the State's academic content standards, such as: o A description of the process used to develop the alternate academic achievement standards that shows: ▪ The State's grade-level academic content standards or extended academic content standards were used as a main reference in writing performance level descriptors for the alternate academic achievement standards; ▪ The process of setting cut scores used, as a main reference, performance level descriptors linked to the State's grade-level academic content standards or extended academic content standards; ▪ The cut scores were set and performance level descriptors written to link to the State's grade-level academic content standards or extended academic content standards; ▪ A description of steps taken to vertically articulate the alternate academic achievement standards (including cut scores and performance level descriptors) across grades.

Critical Element 6.4 – Reporting

	Examples of Evidence
The State reports its assessment results, and the reporting facilitates timely, appropriate, credible, and defensible interpretations and uses of results for students tested by parents, educators,	Collectively, for the State's assessment system, evidence to support this critical element must demonstrate that the State's reporting system facilitates timely, appropriate, credible, and defensible interpretation and use of its assessment results. Evidence to support this critical element both the State's general assessments and AA-AAAS includes:

State officials, policymakers and other stakeholders, and the public, including: • The State reports to the public its assessment results on student achievement at each proficiency level and the percentage of students not tested for all students and each student group after each test administration; • The State reports assessment results, including itemized score analyses, to districts and schools so that parents, teachers, principals, and administrators can interpret the results and address the specific academic needs of students, and the State also provides interpretive guides to support appropriate uses of the assessment results; • The State provides for the production and delivery of individual student interpretive, descriptive, and diagnostic reports after each administration of its assessments that: o Provide valid and reliable information regarding a student's achievement; o Report the student's achievement in terms of the State's grade-level academic achievement standards (including performance-level descriptors); o Provide information to help parents, teachers, and principals interpret the test results and address the specific academic needs of students; o Are available in alternate formats (e.g., Braille or large print) upon request and, to the extent	• Evidence that the State reports to the public its assessment results on student achievement at each proficiency level and the percentage of students not tested for all students and each student group after each test administration, such as: o State report(s) of assessment results; o Appropriate interpretive guidance provided in or with the State report(s) that addresses appropriate uses and limitations of the data (e.g., when comparisons across student groups of different sizes are and are not appropriate). • Evidence that the State reports results for use in instruction, such as: o Instructions for districts, schools, and teachers for access to assessment results, such as an electronic database of results; o Examples of reports of assessment results at the classroom, school, district and State levels provided to teachers, principals, and administrators that include itemized score analyses, results according to proficiency levels, performance level descriptors, and, as appropriate, other analyses that go beyond the total score (e.g., analysis of results by strand); o Instructions for teachers, principals and administrators on the appropriate interpretations and uses of results for students tested that include: the purpose and content of the assessments; guidance for interpreting the results; appropriate uses and limitations of the data; and information to allow use of the assessment results appropriately for addressing the specific academic needs of students, student groups, schools and districts. o Timeline that shows results are reported to districts, schools, and teachers in time to allow for the use of the results in planning for the following school year. • Evidence to support this critical element for both general assessments and AA-AAAS, such as: o Templates or sample individual student reports for each content area and grade level for reporting student performance that: ▪ Report on student achievement according to the domains and subdomains defined in the State's academic content standards and the achievement levels for the student scores (though sub-scores should only be reported when they are based on a sufficient number of items or score points to provide valid and reliable results); ▪ Report on the student's achievement in terms of grade-level achievement using the State's grade-level academic achievement standards and corresponding performance level descriptors; ▪ Display information in a uniform format and use simple language that is free of jargon and understandable to parents, teachers, and principals; ▪ Examples of the interpretive guidance that accompanies individual student reports, either integrated with the report or a separate page(s), including cautions related to the reliability of the reported scores; ▪ Samples of individual student reports in other languages and/or in alternative formats, as applicable. • Evidence that the State follows a process and timeline for delivering individual student reports, such as: o Timeline adhering to the need for the prompt release of assessment results that shows when individual

practicable, in a native language that parents can understand; • The State follows a process and timeline for delivering individual student reports to parents, teachers, and principals as soon as practicable after each test administration.	student reports are delivered to districts and schools; ○ Key documents, such as a cover memo that accompanies individual student reports delivered to districts and schools, listserv messages to district and school test coordinators, or other meaningful communication to districts and schools that include the expectation that individual student reports be delivered to teachers and principals and corresponding expectations for timely delivery to parents (e.g., within 30 days of receipt). Note: Samples of individual student reports and any other sample reports should be redacted to protect personally identifiable information, as appropriate, or populated with information about a fictitious student for illustrative purposes.