



# PEER REVIEW OF STATE ASSESSMENT SYSTEMS



# Overview of Peer Review of State Assessment Systems

October 8, 2015

# AGENDA

1. Purpose and role of peer review
2. Understanding the process
3. Timeline for peer review
4. Understanding the critical elements
  - a. What's new?
  - b. Map of critical elements
  - c. How to read the critical elements
  - d. Introduction to the sections
5. Future supports for States
6. Questions



# PURPOSE AND ROLE OF PEER REVIEW

- States have been required to have statewide assessment systems since the 1994 reauthorization of the ESEA
- Section 1111(b)(3) of the ESEA lays out the requirements of state assessments:
  - The same annual assessments used to measure the achievement of all students in reading/language arts and mathematics in grades 3-8 and once in high school and in science at least once in grades 3-5, 6-9, and 10-12
  - Aligned with the full range of the State's academic content standards
  - Valid, reliable, and consistent with relevant, nationally recognized professional and technical standards for the purposes for which they are used



# PURPOSE AND ROLE OF PEER REVIEW

- Section 1111(e) of the ESEA requires that ED conduct a peer review of State plans, including State assessment systems
  - ED does not review or approve a State's academic content standards
- ED started conducting peer review of assessment systems in 2000
- ED released revised non-regulatory guidance for the peer review of State assessment systems in 2004 to reflect changes in the ESEA in 2002
- Between 2005 and 2012, all States went through peer review for their assessment systems



# PURPOSE AND ROLE OF PEER REVIEW

- In 2012, in light of significant changes in the field, ED paused peer review
  - The AERA/APA/NCME *Standards for Educational and Psychological Testing* were being updated (the revised version was released in 2014)
  - Most States were in the midst of implementing newly adopted content standards and developing new assessment systems
- ED wanted to revise our guidance to reflect lessons learned over the past dozen years, changes in the field and current best practices, revised professional standards, and feedback received from States, experts, and other stakeholders

# PURPOSE AND ROLE OF PEER REVIEW

- The purpose of ED's assessment peer review guidance is three-fold:
  - To support States by identifying expectations that they can use as they develop, administer, and improve their assessment systems in order that they provide valid and reliable information on how well students are achieving a State's challenging academic standards
  - To help States prepare for the peer review of their assessment systems
  - To guide the peer reviewers in the review of the State assessment systems

# UNDERSTANDING THE PROCESS

## WHAT IS REVIEWED?

- The assessment peer review process is:
  - Evidence-based – the peer review is, by nature, backward-looking in order to confirm the technical quality of the assessments based on full administration of the assessments
  - Focused on two primary aspects:
    - Documentation of the process used to develop and administer the assessments
    - Data to confirm the quality of the system (i.e., did the system operate as intended?)



# UNDERSTANDING THE PROCESS

## WHO ARE THE PEER REVIEWERS?

- Assessment peer review is conducted by external assessment experts, including nationally recognized assessment experts, State and local assessment directors, and educators
- ED will soon put out a call for individuals to serve as peer reviewers
- From the full list of peers, ED will develop a small panel of 3-5 peers for each State to review that State's evidence
- ED will conduct an introductory training on the assessment peer review process and criteria prior to beginning the peer review and again before each subsequent review

# UNDERSTANDING THE PROCESS

## HOW DOES THE REVIEW OPERATE?

- Each peer reviewer uses professional judgement to develop individual recommendations to ED regarding whether the documents are sufficient to address the critical element
- ED facilitates a meeting of the peer reviewers to review and discuss the State's documents
- Shortly following the review, ED will provide the State with the notes from the peer reviewers
  - Information and technical assistance to the State
  - Suggest best practices for the State to consider
- Following ED's review and decision, ED will provide formal feedback to the State



# UNDERSTANDING THE PROCESS

## WHO SHOULD SUBMIT FOR REVIEW?

- Almost every State has developed new academic content standards and assessments since ED paused peer review in 2012
- As a result, we believe all States will need to submit documents for peer review for reading/language arts, mathematics, and science
- Includes all tests needed to meet section 1111(b)(3) of the ESEA
  - General assessments
  - Alternate assessments based on alternate academic achievement standards (for students with the most significant cognitive disabilities), and
  - Other assessments (e.g., native language translation, etc.)



# UNDERSTANDING THE PROCESS

## WHO SHOULD SUBMIT FOR REVIEW?

- Some States will be administering new assessments in reading/language arts and/or mathematics in 2015-2016
  - It does not make sense for the State to submit documents about the assessments it administered in 2014-2015
- We believe all States have revised their science assessments since the last time they were peer reviewed
  - We know that many States have recently adopted new science content standards and are developing new science assessments
  - If that is the case, let your program officer know and indicate the timeline for the new science assessments



# TIMELINE FOR REVIEW

- **November 18, 2015** Deadline for States to indicate to their program officer when they will submit documents for peer review
- **Window 1:** Tentatively scheduled for Jan. 25-29, 2016
  - State will need to submit documents by **January 11, 2016**
- **Window 2:** Tentatively scheduled for Mar. 28 – Apr. 1, 2016
  - State will need to submit documents by **March 14, 2016**
- **Window 3:** Tentatively scheduled for May 16-20, 2016
  - State will need to submit documents by **May 2, 2016**





# Understanding the Critical Elements

# WHAT'S NEW?

- General updates – The guidance is revised to follow the new AERA/APA/NCME standards and to reflect changes in assessments over the past 10 years, such as the increased prevalence of technology
- Test security – The guidance includes a larger section focusing on test security before, during, and after the assessments are administered and the State's process to protect the integrity of assessment-related data
- Alternate assessments based on alternate academic achievement standards – The guidance includes much greater detail on evidence related to alternate assessments

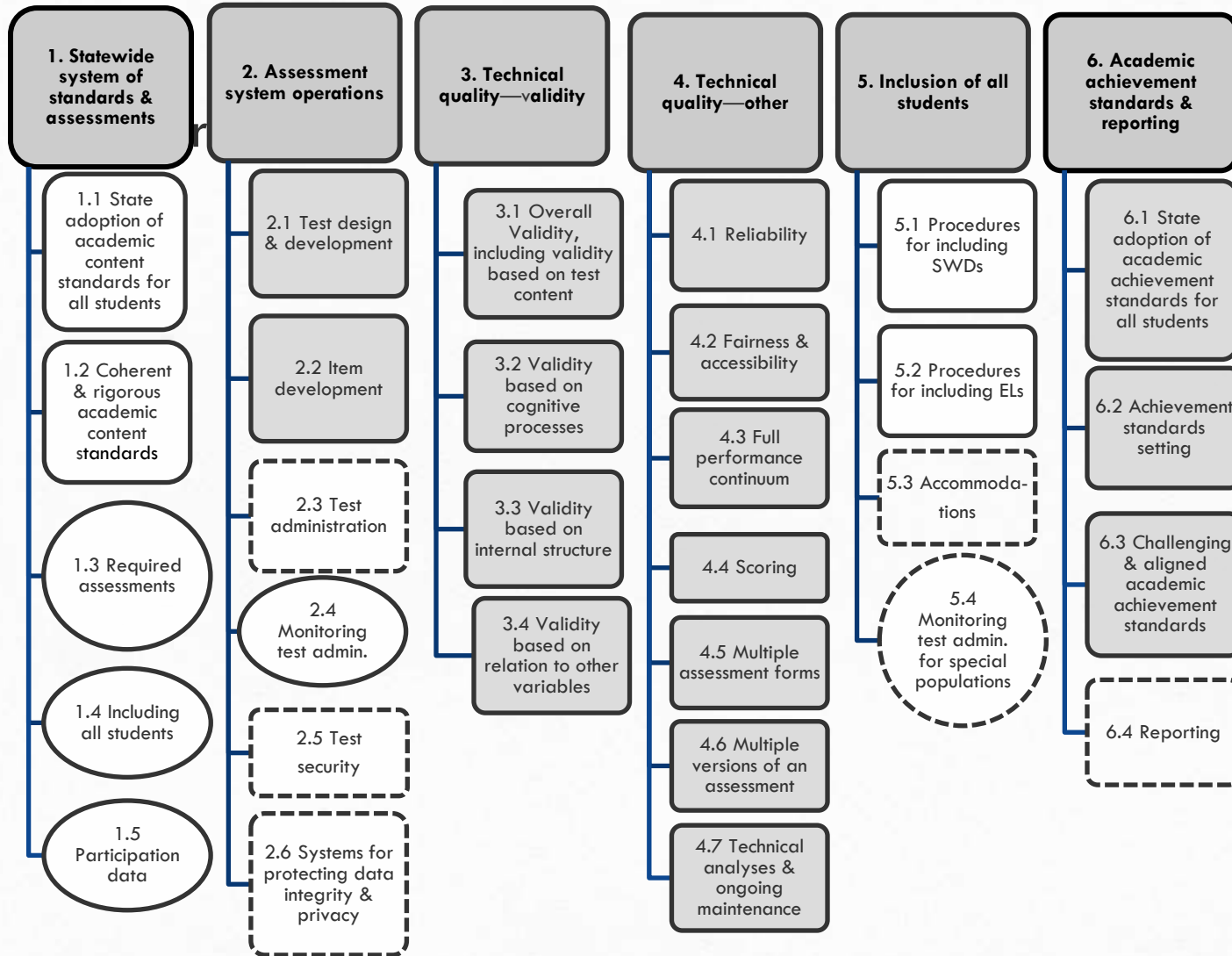
# WHAT'S NEW?

- Coordinated work across States – In cases where States are administering the same assessments, the revised peer review process reduces burden for individual States and ensures consistency by reviewing those submissions together
- Alignment – This is not a new aspect of the review but worth emphasizing—the assessment system is required to cover the full range of the State’s academic content standards
  - For example, if a State has writing as part of its reading/language arts standards, the assessments should include writing
  - Speaking and listening – ED recognizes that large-scale, Statewide assessments may not be ready to include speaking and listening at this time and invites States that have speaking and listening as part of their standards to submit a waiver for these standards, provided the State is working to include them in the future. More details on this will be forthcoming

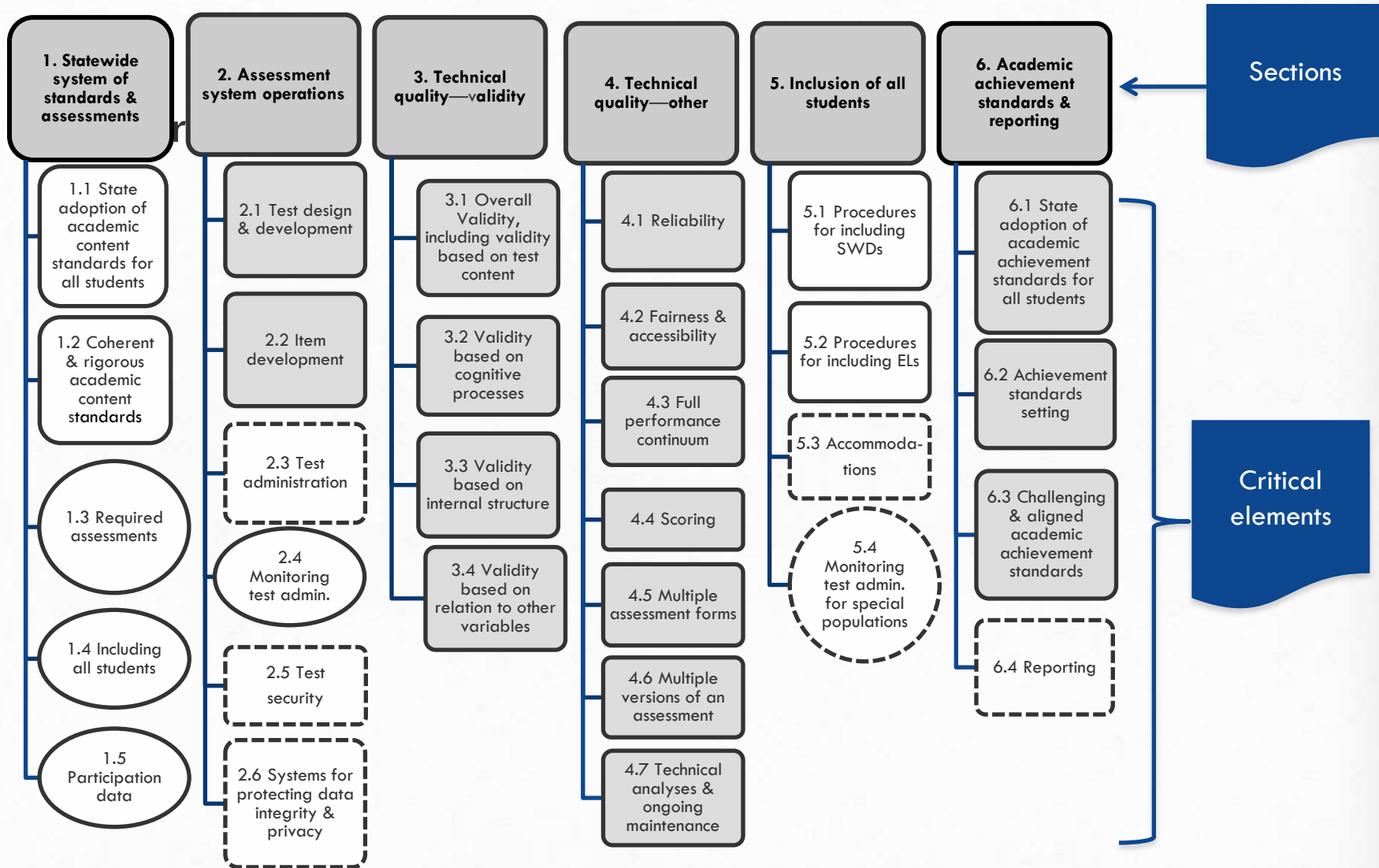




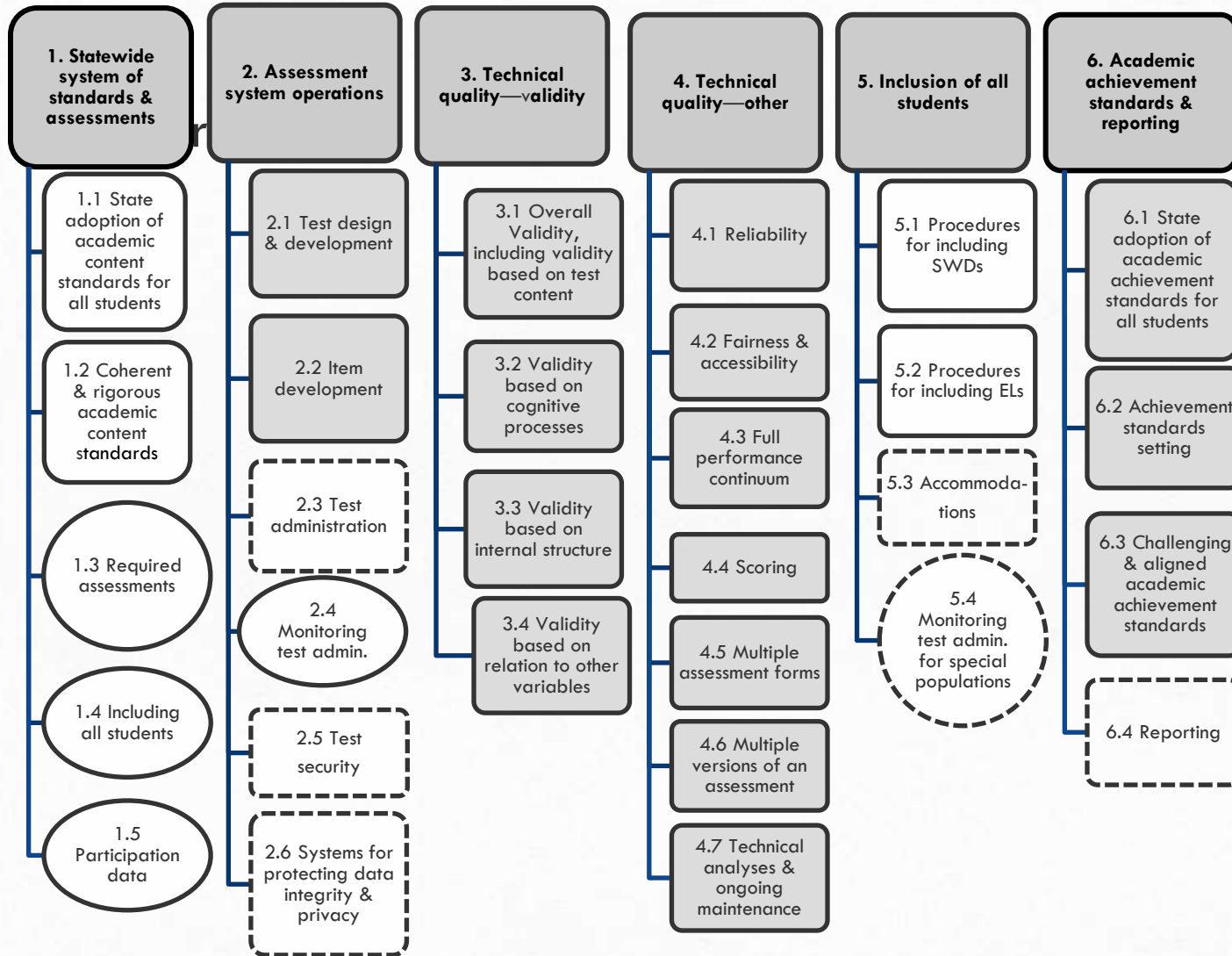
# MAP OF CRITICAL ELEMENTS



# MAP OF CRITICAL ELEMENTS



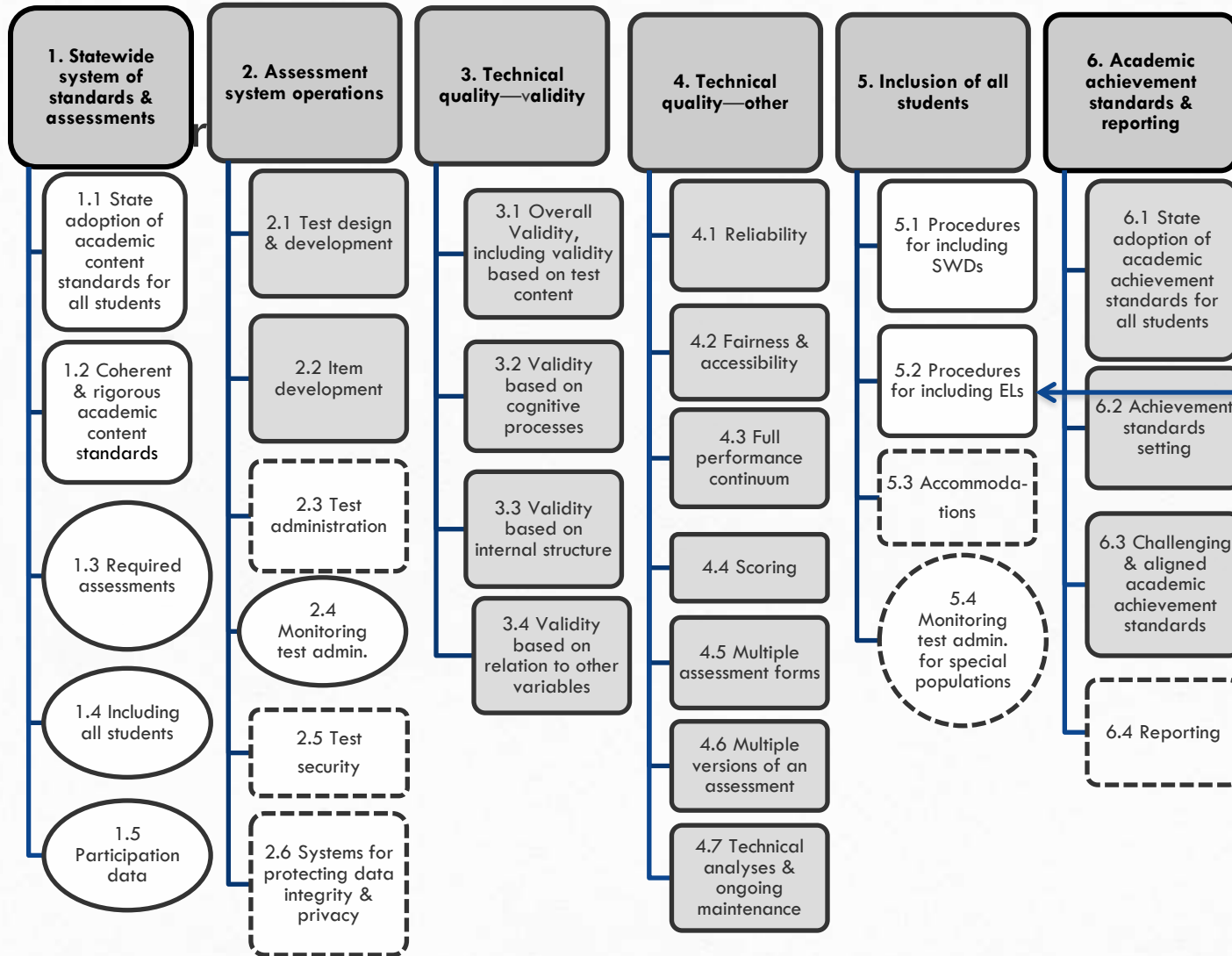
# MAP OF CRITICAL ELEMENTS



Shaded boxes likely addressed by coordinated evidence



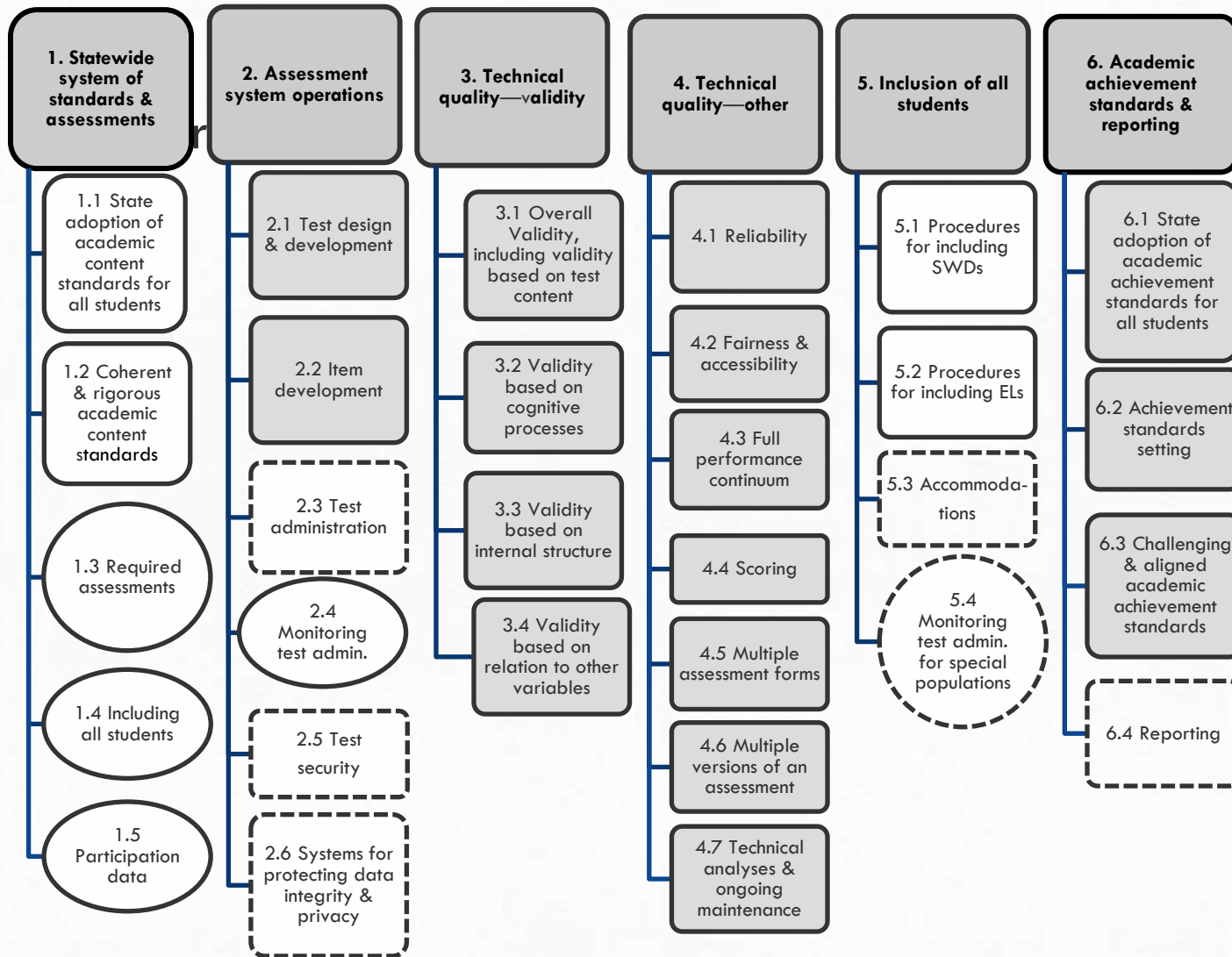
# MAP OF CRITICAL ELEMENTS



Clear boxes likely addressed by State specific evidence for all States



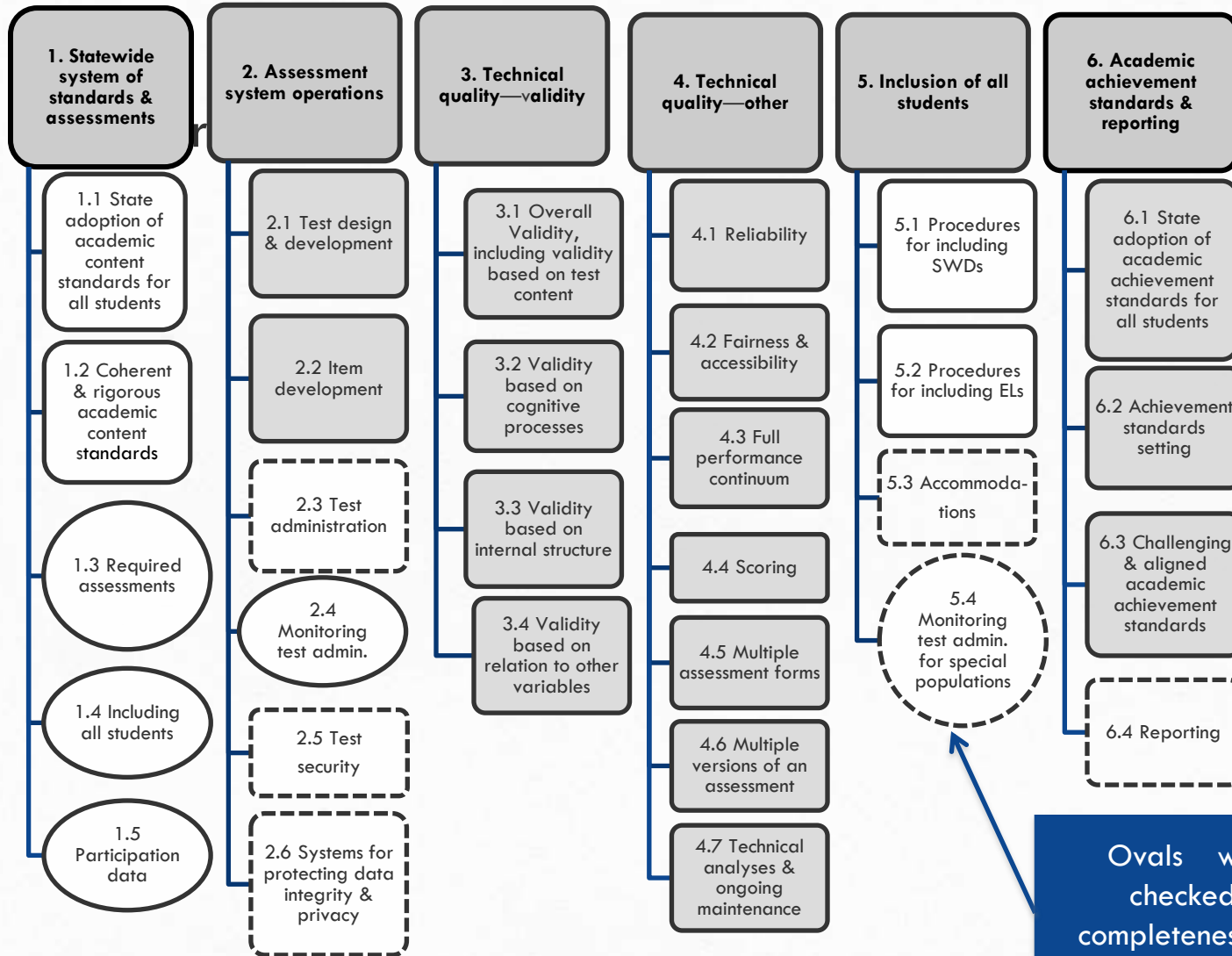
# MAP OF CRITICAL ELEMENTS



Dashed lines likely addressed by a mix of State & coordinated evidence



# MAP OF CRITICAL ELEMENTS



# HOW TO READ THE CRITICAL ELEMENTS

## Critical Element 2.2 – Item Development

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.

Element: State must meet this level of quality

### Examples of Evidence

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 the 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.

**11-115.** 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.



# HOW TO READ THE CRITICAL ELEMENTS

## Critical Element 2.2 – Item Development

	Examples of Evidence
<p>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.</p>	<p>Evidence to support this critical element for the State’s general assessments and AA-AAAS includes documents such as:</p> <p>For the State’s general assessments, evidence, such as a sections in the technical report for the assessments, that show:</p> <ul style="list-style-type: none"> <li>• 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;</li> <li>• A description of the process the State uses to ensure that the 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);</li> <li>• 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;</li> <li>• Description or examples of instructions provided to item writers and reviewers;</li> <li>• 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;</li> <li>• 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;</li> <li>• 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;</li> <li>• 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;</li> <li>• Evidence that the items elicit the intended response processes, such as cognitive labs or interaction studies.</li> </ul> <p><b>11-115.</b> For the State’s AA-AAAS, in addition to the above:</p> <ul style="list-style-type: none"> <li>• 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;</li> <li>• 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.</li> </ul> <p> For the State’s technology-based general assessments and AA-AAAS:</p> <ul style="list-style-type: none"> <li>• Documentation that procedures to evaluate and select items considered the deliverability of the items (e.g., usability studies).</li> </ul> <p>Note: This critical element is closely related to Critical Element 4.2 – Fairness and Accessibility.</p>


Examples of evidence not exhaustive





# HOW TO READ THE CRITICAL ELEMENTS

## Critical Element 2.2 – Item Development


	Examples of Evidence
<p>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.</p>	<p>Evidence to support this critical element for the State’s general assessments and AA-AAAS includes documents such as:</p> <p>For the State’s general assessments, evidence, such as a sections in the technical report for the assessments, that show:</p> <ul style="list-style-type: none"> <li>• 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;</li> <li>• A description of the process the State uses to ensure that the 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);</li> <li>• 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;</li> <li>• Description or examples of instructions provided to item writers and reviewers;</li> <li>• 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;</li> <li>• 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;</li> <li>• 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;</li> <li>• 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;</li> <li>• Evidence that the items elicit the intended response processes, such as cognitive labs or interaction studies.</li> </ul> <p><b>11-115.</b> For the State’s AA-AAAS, in addition to the above:</p> <ul style="list-style-type: none"> <li>• 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;</li> <li>• 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.</li> </ul> <p> For the State’s technology-based general assessments and AA-AAAS:</p> <ul style="list-style-type: none"> <li>• Documentation that procedures to evaluate and select items considered the deliverability of the items (e.g., usability studies).</li> </ul> <p>Note: This critical element is closely related to Critical Element 4.2 – Fairness and Accessibility.</p>

Examples of evidence specific to AA AAAS



# HOW TO READ THE CRITICAL ELEMENTS

## Critical Element 2.2 – Item Development


	Examples of Evidence
<p>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.</p>	<p>Evidence to support this critical element for the State’s general assessments and AA-AAAS includes documents such as:</p> <p>For the State’s general assessments, evidence, such as a sections in the technical report for the assessments, that show:</p> <ul style="list-style-type: none"> <li>• 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;</li> <li>• A description of the process the State uses to ensure that the 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);</li> <li>• 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;</li> <li>• Description or examples of instructions provided to item writers and reviewers;</li> <li>• 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;</li> <li>• 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;</li> <li>• 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;</li> <li>• 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;</li> <li>• Evidence that the items elicit the intended response processes, such as cognitive labs or interaction studies.</li> </ul> <p><b>11-115.</b> For the State’s AA-AAAS, in addition to the above:</p> <ul style="list-style-type: none"> <li>• 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;</li> <li>• 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.</li> </ul> <p> For the State’s technology-based general assessments and AA-AAAS:</p> <ul style="list-style-type: none"> <li>• Documentation that procedures to evaluate and select items considered the deliverability of the items (e.g., usability studies).</li> </ul> <p>Note: This critical element is closely related to Critical Element 4.2 – Fairness and Accessibility.</p>

Examples of evidence specific to AA AAAS



# HOW TO READ THE CRITICAL ELEMENTS

## Critical Element 2.2 – Item Development

	Examples of Evidence
<p>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.</p>	<p>Evidence to support this critical element for the State’s general assessments and AA-AAAS includes documents such as:</p> <p>For the State’s general assessments, evidence, such as a sections in the technical report for the assessments, that show:</p> <ul style="list-style-type: none"> <li>• 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;</li> <li>• A description of the process the State uses to ensure that the 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);</li> <li>• 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;</li> <li>• Description or examples of instructions provided to item writers and reviewers;</li> <li>• 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;</li> <li>• 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;</li> <li>• 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;</li> <li>• 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;</li> <li>• Evidence that the items elicit the intended response processes, such as cognitive labs or interaction studies.</li> </ul> <p><b>11-115.</b> For the State’s AA-AAAS, in addition to the above:</p> <ul style="list-style-type: none"> <li>• 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;</li> <li>• 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.</li> </ul> <p> For the State’s technology-based general assessments and AA-AAAS:</p> <ul style="list-style-type: none"> <li>• Documentation that procedures to evaluate and select items considered the deliverability of the items (e.g., usability studies).</li> </ul> <p>Note: This critical element is closely related to Critical Element 4.2 – Fairness and Accessibility.</p>

Examples of evidence specific to technology based assessments



# FUTURE SUPPORT FOR STATES

- Call for peer reviewers
- State submission cover sheet and index template
- Additional ED webinars
  - Understanding the elements
  - Multiple States submitting documents for a common assessment
  - Lessons learned from window 1
- Resources: [www.ed.gov/admins/lead/account/saa.html](http://www.ed.gov/admins/lead/account/saa.html)
- Contact: Patrick Rooney, [patrick.rooney@ed.gov](mailto:patrick.rooney@ed.gov)



# FUTURE SUPPORT FOR STATES

- CCSSO technical assistance
  - SCASS meetings, October 27-28 in Atlanta
  - One-day meeting, November 10 in DC
    - One person from each state
    - Supported by The Center for Assessment
  - For questions, contact: Scott Norton,  
[Scott.Norton@ccsso.org](mailto:Scott.Norton@ccsso.org)

# FUTURE SUPPORT FOR STATES

HELP US HELP YOU -- IDENTIFY THE SUPPORTS THAT ARE NEEDED

- What questions do you have?
- What support do you need to understand the peer review process?
- What support do you need to understand the critical elements and examples of evidence?
- What support do you need preparing the submission?

*Please type into the chatbox and send a private message to the host.*





# QUESTIONS?



THANK YOU



U.S. DEPARTMENT  
OF EDUCATION





