UNITED STATES DEPARTMENT OF EDUCATION
OFFICE OF ELEMENTARY AND SECONDARY EDUCATION

June 28, 2021

The Honorable Colt Gill
Director
Oregon Department of Education
255 Capitol Street NE
Salem, OR 97310

Dear Director Gill:

Thank you for your participation in the U.S. Department of Education's (the Department) assessment peer review process under Title I of the Elementary and Secondary Education Act of 1965 (ESEA). I appreciate the efforts of the Oregon Department of Education (ODE) to prepare for the peer review, which occurred in July 2020. Specifically, ODE submitted evidence regarding the general assessments for science in grades 5, 8, and high school.

State assessment systems provide essential information that States, districts, principals and teachers can use to identify the academic needs of students, target resources and supports toward students who need them most, evaluate school and program effectiveness and close achievement gaps among students. A high-quality assessment system also provides useful information to parents about their children's advancement against and achievement of grade-level standards. The Department's peer review of State assessment systems is designed to provide feedback to States to support the development and administration of high-quality assessments.

External peer reviewers and Department staff carefully evaluated ODE's submission and the Department found, based on the evidence received, that the components of your assessment system meet some, but not all, of the statutory and regulatory requirements of sections 1111(b)(1) and (2) of the ESEA. Based on the recommendations from this peer review and our own analysis of the State's submission, I have determined the following:

- General assessments for science in grades 5, 8, and high school (Oregon's version of the Cambium Science) - **Partially meets requirements of the ESEA.**

The components that **partially meet requirements** do not meet a number of the requirements of the statute and regulations and/or the ODE will need to provide substantial additional information to demonstrate it meets the requirements. The Department expects that ODE may not be able to submit all of the required information within one year.

A condition has been placed on ODE's Title I, Part A grant due to the assessment peer review outcome status identified above. The condition will remain until such time as the State satisfies all assessment system requirements.

The specific list of items required for ODE to submit is enclosed with this letter. I request that ODE submit a plan within 30 days outlining when it will submit all required additional documentation for peer review. I recognize the unprecedented situation affecting you and your schools due to widespread and extended school closures caused by the novel coronavirus, COVID-19. As a result, if you need more than 30 days to submit your plan, please let my staff know at ESEA.Assessment@ed.gov. Upon submission of the plan, the Department will reach out to ODE to determine a mutually agreeable schedule. Resubmission should occur once all necessary evidence is complete (rather than in multiple submissions).

The full peer review notes from the review are enclosed. These recommendations to the Department formed the basis of our determination. Please note that the peers' recommendations may differ from the Department's feedback; we encourage you to read the full peer notes for additional suggestions and recommendations for improving your assessment system beyond what is noted in the Department's feedback. Department staff will reach out to your assessment director in the next few days to discuss the peer notes and the Department's determination and to answer any questions you have.

Thank you for your ongoing commitment to improving educational outcomes for all students. I look forward to our continued partnership as we move ahead with this critical work. I appreciate the work you are doing to improve your schools and provide a high-quality education for your students.

If you have any questions, please contact my staff at: ESEA.Assessment@ed.gov.

Sincerely,

/s/
Ian Rosenblum
Deputy Assistant Secretary for Policy and Programs
Delegated the Authority to Perform the
Functions and Duties of the Assistant Secretary
Office of Elementary and Secondary Education

Enclosures

cc: Dan Farley, ODE Director of Assessment

Critical Elements that Require Additional Evidence for Oregon's Assessment System

Critical Element	Additional Evidence Needed
2.1 – Test Design and Development	For Oregon's version of Cambium science: • Evidence that the State's test design and test development process includes statement(s) of the purposes of the assessments and the intended interpretations and uses of results (e.g., evidence of an explicit rather than implicit description of the purposes and interpretations of the uses of assessment results). • Evidence that the State's test design and test development process includes test blueprints that describe the structure of each assessment in sufficient detail to support the development of assessments that are technically sound, measure the depth and breadth of the State's grade-level academic content standards, and support the intended interpretations and uses of the results, specifically: o Evidence that guidelines given to participating States on developing test blueprints include specific rules, constraints, and parameters for building test blueprints; and evidence of the procedure for reviewing each State's test blueprints. o Evidence that State-developed test blueprints consider the full range (all three dimensions) of the science standards rather than just the disciplinary core ideas dimension. o Evidence that shows the full range (all three dimensions) of the science standards play a role throughout the test design and development process such as with test blueprints, scoring, and results; or give a rationale for why the test blueprints will not consider all three science dimensions throughout the test design and development process
2.2 – Item Development	For Oregon's version of Cambium science: • Evidence that 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, specifically: o Evidence that clarifies discrepancies between the expected vocabulary listed in the item specifications and information provided in the science frameworks and Next Generation Science Standards. o Evidence on the demographic characteristics of fairness and other committee review members for each participating State, especially in terms of their science expertise.
2.3 – Test Administration	For Oregon's version of Cambium science: • Evidence that the State has established contingency plans to address possible technology challenges during test administration (e.g., evidence at the vendor level).

Critical Element	Additional Evidence Needed
	• Evidence that the State has established procedures to ensure that general and special education teachers, paraprofessionals, teachers of English learners, specialized instructional support personnel, and other appropriate staff receive necessary training to administer assessments, including, as necessary, alternate assessments, and know how to make use of appropriate accommodations during assessments for all students with disabilities (e.g., evidence of training attendance sheets for school staff rather than just for district test coordinators).
2.4 – Monitoring Test Administration	For Oregon's version of Cambium science: • Evidence that the State adequately monitors the test administration of its State assessments to ensure that standardized test administration procedures are implemented with fidelity across districts and schools (e.g., evidence of scheduled visits with districts or schools, or redacted letters sent to districts or schools following the in-person monitoring visits).
2.5 – Test Security	For Oregon's version of Cambium science: • Evidence that the State has implemented and documented an appropriate set of policies and procedures to prevent test irregularities and ensure the integrity of test results, specifically: o Evidence of a process for remediation following test security incidents that occur during test development and administration under the consortium's control, and a process for communicating this information to participating States in a timely manner. o Evidence of State policies and procedures to prevent and detect test irregularities (e.g., evidence of a clearer description of the vendor's test security process for Oregon).
2.6 – Systems for Protecting Data Integrity and Privacy	For Oregon's version of Cambium science: • Evidence that 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 (e.g., evidence of provisions for ensuring that each participating State's data is separated from other State's data).
3.1 – Overall Validity, including Validity Based on Content	For Oregon's version of Cambium science: • Evidence that the State's academic assessments measure the knowledge and skills specified in the State's academic content standards (e.g., evidence of a plan to address issues identified in the shared item bank alignment report such as the editorial errors and cases where items included expectations beyond grade-level).
3.2 – Validity Based om Cognitive Processes	For Oregon's version of Cambium science: • Evidence that 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 (e.g., evidence of demographic characteristics for panelists

Critical Element	Additional Evidence Needed
	involved in the State's shared item bank alignment study, especially in terms of their science expertise).
3.3 – Validity Based on Internal Structure	For Oregon's version of Cambium science: • Evidence the State has documented adequate validity evidence that the scoring and reporting structures of the assessments are consistent with the subdomain structure of the State's academic content standards (e.g., evidence of a rationale for why the science disciplines of life, physical, and earth and space were chosen as the subdomain structure of the assessments rather than all three dimensions of the science standards).
4.1 – Reliability	For Oregon's version of Cambium science: • Evidence that the State has documented adequate reliability evidence for its assessments for the State's population overall and each student group consistent with nationally recognized professional and technical testing standards, specifically: o Evidence that the overall standard error of measurement (SEM) has been reported. o Evidence of a plan to address the large conditional SEM in the tails of the distributions for each grade level.
4.2 – Fairness and Accessibility	For Oregon's version of Cambium science: • Evidence that the State has taken reasonable and appropriate steps to ensure that its assessments are accessible to all students and fair across student groups in their design, development, and analysis (e.g., evidence that the formatting and technology-related issues observed in the Braille cognitive lab study have been resolved).
4.4 – Scoring	For Oregon's version of Cambium science: • Evidence that the State has established and documented standardized scoring procedures and protocols for its assessments that are designed to produce reliable and meaningful results, facilitate valid score interpretations, and report assessment results in terms of the State's academic achievement standards (e.g., evidence on how paper test forms are scored).
4.6 – Multiple Versions of an Assessment	For Oregon's version of Cambium science: • Evidence that the State followed a design and development process to support comparable interpretations of results for students tested across the paper, Braille, and online versions of the assessments; and the Spanish and English versions.
4.7 – Technical Analysis and Ongoing Maintenance	For the Cambium science: • Evidence that the State has a system for monitoring, maintaining, and improving, as needed, the quality of its assessment system, including clear and technically sound criteria for the analyses of all the assessments in its assessment system, specifically: o Evidence to clarify the details for continued item development and replenishment of the item bank. o Evidence that adequate technical quality is made public, including on the State's website.

Critical Element	Additional Evidence Needed
5.3 – Accommodations	For Oregon's version of Cambium science: • Evidence that accommodations the State provides are appropriate and effective for meeting the individual student's need(s) to participate in the assessments, do not alter the construct being measured, allow meaningful interpretations of results and comparison of scores for students who need and receive accommodations and students who do not need and do not receive accommodations, specifically: o Evidence of a rationale for not adding the bilingual/dual language word-to-word dictionary to the list of allowable supports given that the consortium allows for its use. o Evidence from literature reviews or professional organizations that the accommodations provided to students with disabilities and English learners allow for valid inferences.
5.4 – Monitoring Test Administration for Special Populations	For Oregon's version of Cambium science: • Evidence that the State monitors test administration in districts and schools to ensure that appropriate assessments, with or without accommodations, are selected for all students with disabilities and English learners (e.g., evidence of monitoring test administration for special populations such as with redacted letters to districts and schools).
6.2 – Achievement Standards-Setting	For Oregon's version of Cambium science: • Evidence that the State used a technically sound method and process that involved panelists with appropriate experience and expertise for setting academic achievement standards (e.g., evidence on the science expertise of the standards-setting panel).
6.3 – Challenging and Aligned Academic Achievement Standards	For Oregon's version of Cambium science: • Evidence that the State's academic achievement standards are aligned with entrance requirements for credit-bearing coursework in the system of public higher education in the State and relevant State career and technical education standards (e.g., evidence of comparisons between Cambium science scores on the high school assessments and scores on career ready tests like WorkKeys).
6.4 – Reporting	For Oregon's version of Cambium science: • Evidence that the State reports its assessment results for all students assessed, and the reporting facilitates timely, appropriate, credible, and defensible interpretations and uses of those results by parents, educators, State officials, policymakers and other stakeholders, and the public, specifically: o Evidence that assessment reports are to the extent practicable written in a language that parents and guardians can understand or, if it is not practicable to provide written translations to a parent or guardian with limited English proficiency, are orally translated for such parent or guardian. o Upon request by a parent who is an individual with a disability as defined by the American with Disabilities Act (ADA), as

Critical Element	Additional Evidence Needed
	amended, are provided in an alternative format accessible to that parent.

U. S. DEPARTMENT OF EDUCATION

Peer Review of State Assessment Systems

July-August 2020 State Assessment Peer Review Notes

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

Note: Peer review notes provide the combined recommendations of the individual peers to the U.S. Department of Education (Department), based on the statute and regulations, the Department's peer review guidance, and the peers' professional judgement of the evidence submitted by the State. These assessment peer review notes, however, do not necessarily reflect 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 for assessment peer review. Although the peer notes inform the Secretary's consideration of each State's assessment system, the Department makes the final decision regarding whether the assessment system meets the requirements in the statute and regulations. As a result, these peer notes may not completely align with the final determination made by the Department.

Contents

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

SECTION 1: STATEWIDE SYSTEM OF STANDARDS AND ASSESSMENTS

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

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
For academic content standards: 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 schools and public school students in the State.	**CE 1.1** • *_Evidence#OR1.1.1_OregonStateBoardAuthority* • *_Evidence#OR1.1.2_March6-2014-board-agenda* • *_Evidence#OR1.1.3_March-6-7-2014-draft-minutes* • *_Evidence#OR1.1.4_March6-2014-board-actions* • Evidence#OR2.1.1.2_OAR581_022_2270 (added by peers from submission from another CE)	*Peers found evidence that the State formally adopted challenging academic content standards for all students in science and applies its academic content standards to all public schools and public school students in the State.* Requirement met based on evidence provided by state standards – state adopted the Next Generation Science Standards (NGSS) (OR1.1.4), and provided evidence of SBE's authority to apply the standards to public schools & students (OR.1.1.1). OR 2.1.1.2 (p. 1) under Division 22 of Standards for Public Elementary and Secondary Schools states that "Each school district shall assess and record each student's progress and achievement in all subject areas of instruction and to academic content standards consistent with ORS 329.045 and OAR 581-022-2030."

Section 1.1 Summary Statement

_X__ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:
- [list additional evidence needed w/brief rationale]

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 1.2 – Challenging Academic Content Standards

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
For academic content standards: The State's challenging academic content standards in reading/language arts, mathematics, and science are aligned with entrance requirements for credit-bearing coursework in the system of public higher education in the State and relevant State career and technical education standards.	**CE 1.2** • *_Evidence#OR1.2.1_orscistand-ngss-k-12_2014* • *_Evidence#OR1.2.2_Developing the Standards _ Next Generation Science Standards* • *_Evidence#OR1.2.3_Summary of Science Panel Recommendations to Present to the SBE* • *_Evidence#OR1.2.4_K_12_Framework* • *_Evidence#OR1.2.5_2014 Oregon Science Content and Asmt Panel Membership* • *Evidence#OR1.1.3_March-6-7-2014-draft-minutes*	*Peers found evidence that representatives of OR universities were included in the panel that reviewed and recommended state adoption of NGSS (OR.1.2.5). Evidence indicates that Oregon adopted the Next Generation Science Standards (NGSS) that are designed to "prepare students for college and careers" (OR1.2.2, p. 2).* *Peers did **not** find evidence of alignment between NGSS and requirements for credit bearing coursework or CTE standards. Peers did not find evidence of the process followed by the panel (OR.1.2.5) that reviewed the NGSS to ensure alignment with credit bearing coursework or CTE standards.* Examples of evidence that could address this CE include an alignment study that would go beyond a statistical validity/reliability study, in that it would evaluate the conceptual linkages between: - NGSS/OSAS and requirements for higher education credit bearing coursework - NGSS/OSAS and CTE standards.

Section 1.2 Summary Statement

___ No additional evidence is required or

_X__ The following additional evidence is needed/provide brief rationale:

- Evidence of alignment of academic content standards with entrance requirements for credit-bearing coursework in the system of public higher education in the State and relevant State career and technical education standards.

Critical Element 1.3 – Required Assessments

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State's assessment system includes annual general and alternate assessments aligned with **grade-level academic achievement standards** or alternate academic achievement standards in: • Reading/language arts (R/LA) and mathematics in each of grades 3-8 and at least once in high school (grades 9-12); • Science at least once in each of three grade spans (3-5, 6-9 and 10-12). AND The State's **academic content assessments** must be the same assessments administered to all students in the tested grades, with the following exceptions: • Students with the most significant cognitive disabilities may take an alternate assessment aligned with alternate academic achievement standards. • A State may permit an LEA to administer a nationally recognized high school academic assessment in lieu of the State high school assessment if certain conditions are met. • A State that administers an end-of-course high school mathematics assessment may exempt an 8th grade student from the mathematics assessment typically administered in	Reviewed by Department Staff Only	Department staff determined that the State's evidence is sufficient for this critical element.

eighth grade and allow the student to take the State end-of-course mathematics test instead. • The Department may have approved the State, under the Innovative Assessment Demonstration Authority, to permit students in some LEAs to participate in a demonstration assessment system in lieu of participating in the State assessment.		
Section 1.3 Summary Statement		
_X__ No additional evidence is required or ___ The following additional evidence is needed/provide brief rationale:		

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 1.4 – Policies for Including All Students in Assessments

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or 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 those children 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 ELs: o Policies state that all ELs must be included in all aspects of the content assessment system, unless the State has chosen the statutory option for recently arrived ELs under which such ELs are exempt from one administration of its reading/ language arts assessment. o If a State has developed native language assessments for ELs in R/LA, ELs must be assessed in R/LA 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	Reviewed by Department Staff Only	Department staff determined that the State's evidence is sufficient for this critical element.

language assessments for a period not to exceed two additional consecutive years. ○ If the State uses the flexibility for Native American language schools and programs: (1) the State provides the content assessment in the Native American language to all students in the school or program; (2) the State submits such content assessment for peer review as part of its State assessment system; and (3) the State continues to provide ELP assessments and services for ELs as required by law. The State must assess in English the students' achievement in R/LA in high school.		

Section 1.4 Summary Statement

_X__ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:

- [list additional evidence needed w/brief rationale]

Critical Element 1.5 – Meaningful Consultation in the Development of Challenging State Standards and Assessments

(Note: this is a new requirement under ESSA, so it does not apply to standards and assessments adopted prior to the passage of ESSA (December 2015)).

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
If the State has developed or amended challenging **academic** standards and assessments, the State has conducted meaningful and timely consultation with: • State leaders, including the Governor, members of the State legislature and State board of education (if the State has a State board of education). • Local educational agencies (including those located in rural areas). • Representatives of Indian tribes located in the State. • Teachers, principals, other school leaders, charter school leaders (if the State has charter schools), specialized instructional support personnel, paraprofessionals, administrators, other staff, and parents.	Reviewed by Department Staff Only	Oregon adopted the Next Generation Science Standards (NGSS) on March 6, 2014 which is prior to the passage of ESSA (December 2015).

Section 1.5 Summary Statement

_X__ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:

- [list additional evidence needed w/brief rationale]

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

SECTION 2: ASSESSMENT SYSTEM OPERATIONS

Critical Element 2.1 – Test Design and Development

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State's test design and test development process is well-suited for the content, is technically sound, aligns the assessments to **the depth and breadth of the State's academic content standards** for the grade that is being assessed and includes: • 2.1.1 Statement(s) of the purposes of the assessments and the intended interpretations and uses of results;	**CE 2.1** • *_Please see Common Peer Review Submission for Science Supplemental Evidence and Notes* ***2.1.1*** • *_Evidence#OR2.1.1.1_V1_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR1.3.1_2018_19_TAM* • *_Evidence#OR1.3.2_OAR581_022_2100* • *_Evidence#OR2.1.1.2_OAR581_022_2270* • *_Evidence#OR2.1.1.3_OAR581_022_2250* • *_Evidence#OR2.1.1.4_ ODE Assessment Homepage 05_20_2020* • *Evidence#OR2.1.1.5_ ODE Understanding Student Assessment _webpage_ 05_20_2020* • *_Evidence#OR2.1.1.6_ ODE Statewide Assessment Overview _webpage 05_20_2020*	**CE 2.1.1:** *Peers found evidence of statement(s) of the purposes of the assessments and the intended interpretations and uses of results.* Oregon provided a list of intended uses (OR2.1.1.1, pgs. 2-3) that is very similar to what the Common submission provided. This list starts with "indicator of academic achievement and progress aligned to academic content standards," which is certainly a primary use of the assessment. However, the rest of the list has bulleted elements that are simply actions related to test administration, scoring, and reporting, without regard to actual uses of the results. This list is derived from the relevant sections of their State code, the Oregon Revised Statute (ORS) 329.485, and from the federal Every Student Succeeds Act (ESSA) plan. Oregon also included in its description of purposes to "support instruction and student learning by providing valuable feedback to educators and parents which can be used to form instructional strategies to remediate or enrich instruction. … and to monitor improvement at the student and group levels over time." These are important uses of the results and imply that the results will be interpreted in terms of student learning and progress. *OR2.1.1.5* from the ODE website lists additional uses as: • Teachers and administrators use summative assessment results to review learning patterns

		annually, to determine systems-level changes that might be required from year-to-year; • State and local leaders use summative assessment results to make important policy decisions…"
• 2.1.2 Test blueprints that describe the structure of each assessment in sufficient detail to support the development of assessments that are technically sound, measure the depth and breadth of **the State's grade-level academic content standards** and support the intended interpretations and uses of the results.	***2.1.2*** *• _Evidence#OR1.5.4.1_ V2_OSAS Science Assessment Tech Report* *• _Evidence#OR2.1.2.2_Oregon Science Assessment Blueprint* *• _Evidence#OR2.1.2.3_ Oregon Science Assessment Blueprint Summary* *• _Evidence# OR2.1.2.4_OSAS Science Assessment Info* *• _Evidence#OR2.1.2.5_Flyer OSAS Science* *• _Evidence#OR2.1.2.6_V4_OSAS Science Tech Report 2018-2019*	**CE 2.1.2:** *Peers did **not** find evidence of test blueprints that describe the structure of each assessment in sufficient detail to support the development of assessments that **measure the depth and breadth of the State's grade-level academic content standards and support the intended interpretations and uses of the results.*** *Peers found evidence of test blueprints that describe the structure of each assessment in sufficient detail to support the development of assessments that are **technically sound**.* It is unclear how the algorithm samples within the grade level PEs (since the assessment is a grade-band, not grade-level assessment), and whether these are done consistently by test form. For example, it appears possible for some students to have a test composed entirely of below-grade-level (i.e. grade 3 & 4) item clusters for the grade 5 assessment. Evidence is sufficient to support that Oregon followed the guidelines provided by Cambium Assessment, Inc. (CAI). As noted in the Common submission, the guidelines provided by CAI in the development of test blueprints lead to blueprints being based on Disciplinary Core Ideas (DCI) (i.e., content) coverage alone, without regard to the other two dimensions: Science and Engineering Practices (SEPs) and Cross-Cutting Concepts (CCCs). To properly meet the breadth and depth of the standards, developers need to provide ways to ensure that participating states have blueprints that allow users to evaluate students' use of the eight SEPs as well as which clusters/items include student use of the CCCs to demonstrate sense-making. Since Oregon's uses of the assessment include instructional feedback to educators and to policy makers with regard to

		learning patterns, it is vital that the blueprints ensure coverage of all three dimensions of the standards.
• 2.1.3 Processes to ensure that each academic 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).	***2.1.3*** • *_Evidence#OR1.5.4.1_V2_OSAS Science Assessment Tech Report 2018-2019* • *A_Oregon2018_19SciencePeerReviewIndex_VF, CE 4.4, pdf pg. 65, pg. 62* (added by peers)	**CE 2.1.3** *Peers found evidence of processes to ensure that each academic assessment is tailored to the knowledge and skills included in* ***the State's academic content standards.*** *Peers did* ***not*** *find evidence of processes to ensure that each academic assessment reflects appropriate inclusion of challenging content and requires complex demonstrations or applications of knowledge and skills (i.e., higher-order thinking skills).* Peers note that constructed response items from the consortium item bank are excluded from the Oregon assessment design (see Oregon Index, CE 4.4, pdf pg. 65, pg. 62), and the independent alignment study identified no items on the assessment forms reviewed as being at Depth of Knowledge (DOK) level 4, thereby indicating a limitation in the extent to which the assessment is measuring higher-order thinking skills. As mentioned in 2.1.2, because only DCI (content) is represented in the blueprints, this leaves open the possibility that the assessment could be overly focused on content. The three-dimensionality of the NGSS framework is designed to ensure that science achievement reflects the ability to make sense of challenging content and to perform complex demonstrations and applications of knowledge and skills. If the emphasis in this assessment is on content (what they know) alone, there is no assurance that students' higher-order thinking skills are being elicited and to demonstrate what they are able to do with that content. It is important for the states within the consortium to seek guidance from CAI on how to ensure that their assessments include a proper balance of all three dimensions.
	2.1.4	

• 2.1.4 If the State administers computer-adaptive assessments, the item pool and item selection procedures adequately support the test design and intended uses and interpretations of results.	*OR1.5.4.1 – V2 OSAS Science Assessment Tech Report (added by peers)* *Evidence#OR2.1.1.1_V1_OSAS Science Assessment Tech Report 2018-2019 (added by peers)*	**CE 2.1.4:** *Peers found evidence that the state's computer-adaptive assessments using the Linear-on-the-Fly (LOFT) design, the item pool and item selection procedures adequately support the test design and intended uses and interpretations of results.* While the design of the Oregon science assessment is not computer adaptive in the traditional sense of students receiving items based on their performance on individual items on the test, it is in fact computer adaptive in the sense that students do not all take the same items. Oregon chose to use a LOFT design, in which the selection of items for each student varies according to parameters designed to ensure that the test blueprint is maintained. The Common submission demonstrated that the processes used in the development and implementation of the LOFT approach appear to adequately support the test design in an overall sense. *OR1.5.4.1 – V2 (*pg. 46-47) describes the results of studies on the blueprint match and item exposure rates for Oregon's 2019 operational assessment, to determine how well the LOFT design worked when applied to their operational test. Both the English and the Spanish online tests in all three grades met the blueprint specifications with a 100% match at all content levels, except for three students out of 44,989 in grade 5 (who may have seen items on a previous administration of the test at the same grade level, which would have limited the item bank available for them). Additional evidence was reported on p. 68 of *Evidence#OR2.1.1.1_V1_OSAS Science Assessment Tech Report 2018-2019* regarding item exposure rates. Because there will often be students who may be repeating a grade level, or who might start a test and then have it reset to start again, it is important for Oregon to work with

		the consortium to ensure that the item bank is replenished and expanded on a regular basis.
• 2.1.5 If the State administers a computer-adaptive assessment, it makes proficiency determinations with respect to the grade in which the student is enrolled and uses that determination for all reporting.	***2.1.5*** See Common submission.	**CE 2.1.5:** *Peers did **not** find sufficient evidence that the computer-adaptive assessment (LOFT design) makes proficiency determinations with respect to the grade in which the student is enrolled and uses that determination for all reporting.* Because the blueprints are built on grade band (not grade level), it is unclear that the LOFT algorithm is able to ensure the test forms are consistently created in terms of distribution of clusters by grade level.
• 2.1.6 If the State administers a content assessment that includes portfolios, such assessment may be partially administered through a portfolio but may not be *entirely* administered through a portfolio.	***2.1.6*** *Not applicable*	**CE 2.1.6:** *Not applicable*

Section 2.1 Summary Statement

___ No additional evidence is required or

_X__ The following additional evidence is needed/provide brief rationale:

- Test blueprints that describe the structure of each assessment in sufficient detail to support the development of assessments that measure the depth and breadth of **the State's grade-level academic content standards** and support the intended interpretations and uses of the results in terms of the three dimensions of the NGSS.
- Evidence of processes to ensure that each academic assessment reflects appropriate inclusion of challenging content and requires complex demonstrations or applications of knowledge and skills (i.e., higher-order thinking skills).
- Evidence that the computer-adaptive assessment (LOFT design) makes proficiency determinations with respect to the grade in which the student is enrolled and uses that determination for all reporting.

Critical Element 2.2 – Item Development

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or 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.	**CE 2.2** • *_Please see Common Peer Review Submission for Science Supplemental Evidence and Notes* • *A_Oregon2018_19SciencePeerReviewIndex_VF, CE 4.4, pdf pg. 65, pg. 62* (added by peers)	**CE 2.2:** *Because Oregon adopted the NGSS without modifications, the item development process documented in the Common submission applies in total to the Oregon assessment.* However, it is unclear as to why Oregon selected items for their state assessment that excluded all constructed-response items. See Oregon State Index, CE 4.4, pdf p. 65, p. 62.

Section 2.2 Summary Statement

_ __ No additional evidence is required or

_X_ The following additional evidence is needed/provide brief rationale:

- Rationale for the State selecting only items from the shared item bank that are not constructed-response items and how their selection of items ensures that the assessment assesses student achievement based on the State's academic content standards in terms of content and cognitive process, including higher-order thinking skills.

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 2.3 – Test Administration

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State implements policies and procedures for standardized test administration; specifically, the State: • 2.3.1: Has established and communicates to educators clear, thorough and consistent standardized procedures for the administration of its assessments, including administration with accommodations;	**CE 2.3** • *_Please see common submission for supplemental evidence and notes.* ***2.3.1*** • *_Evidence#OR1.3.1_2018_19_TAM* • *_Evidence#OR1.4.2_2018-19_OAM_Final* • *_Evidence#OR1.3.2_OAR581_022_2100* • *_Evidence#OR2.3.1.1_DTC Training Materials Webscreen and Training Modules* • *_Evidence#OR2.3.1.2_DTC Training Attendees* • *_Evidence#OR2.3.1.3_Mandatory Training Notification* • *_Evidence#OR2.3.1.4_2018DTC_Webinar_VF* • *_Evidence#OR2.3.1.5_AA UPdate webpage This does not seem to provide any relevant evidence.--peers* • *_Evidence#OR2.3.1.6_OSASPortal Announcements*	**CE 2.3.1:** *Peers found evidence that the state has established and communicates to educators clear, thorough and consistent standardized procedures for the administration of its assessments, including administration with accommodations.* One area of concern is that several instances in the evidence use the phrasing "… highly **recommended** that all students access the Sample Tests and Training Tests" (*Evidence#OR1.3.1_2018_19_TAM, p.44, 45)* To have this procedure be a recommendation rather than a requirement can result in some students not having this exposure. Part of the procedure should include TAs documenting for each student that the student participated actively in taking one or more practice tests.
• 2.3.2: Has established procedures to ensure that general and special education teachers, paraprofessionals, teachers of ELs, specialized instructional support personnel, and other appropriate staff receive necessary training to administer assessments and know how to administer assessments, including, as necessary, alternate assessments, and know how to	***2.3.2*** • *_Evidence#OR1.3.1_2018_19_TAM* • *_Evidence#OR2.3.1.1_DTC Training Materials Webscreen and Training Modules* • *_Evidence#OR2.3.1.4_2018DTC_Webinar_VF* • *_Evidence#OR2.3.1.2_DTC Training Attendees* • *_Evidence#OR2.3.1.3_Mandatory Training Notification*	**CE 2.3.2:** *Peers did **not** find sufficient evidence that the State has established procedures to ensure that general and special education teachers, paraprofessionals, teachers of ELs, specialized instructional support personnel, and other appropriate staff receive necessary training to administer assessments and know how to administer assessments, including, as necessary, alternate assessments, and know how to make use of appropriate accommodations during assessments for all students with disabilities.* While Oregon provides a training quiz for District Test Coordinators (DTCs), as indicated in

make use of appropriate accommodations during assessments for all students with disabilities;		*Evidence#OR2.3.1.4_2018DTC_Webinar_VF* (p. 3), no evidence was found that Test Administrators (TAs) are required to complete a quiz to certify their understanding of proper test administration. Since the Common submission indicates that CAI created an interactive online Certification Course for TAs (with participants starting test sessions under different scenarios and answering multiple-choice questions along the way to solidify their understanding), it would be important for Oregon to clarify why this piece is not included in their TA training. Peers did not find evidence that all school districts/LEAs were included in the webinar DTC training (and what the procedures were for districts that had no representation). On p. 15 of the TAM, it indicates that "TAs who will administer either online assessments through the Braille Interface or the Oregon Extended Assessments must receive additional specialized training from ODE or its designee in addition to receiving the test administration and security training required for all TAs." However, the evidence did not include details about what this training looks like, for readers or scribes, or TAs giving Braille, ASL, or paper-on-demand versions of the assessment. The submission referenced training modules built by Smarter Balanced Assessment Consortium (SBAC). Those modules could have been provided.
• 2.3.3: If the State administers technology-based assessments, the State has defined technology and other related requirements, included technology-based test administration in its	***2.3.3*** • *_Evidence#OR1.3.1_2018_19_TAM* • *_Evidence#OR2.3.3.1_BroganPRLetter_10_30_2019 This does not seem to provide any relevant evidence. It applies to the ELP test, not science. --peers*	**CE 2.3.3**: *Peers found evidence that 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.*

standardized procedures for test administration, and established contingency plans to address possible technology challenges during test administration.	• *_Evidence#OR2.3.3.2_18_19 Technical Specifications Manual* • *_Evidence#OR2.3.3.3_AIRBusinessContinuityPlan_PROPRIETARY* • *_Evidence#OR2.3.3.4_HelpDesk Contract* • *_Evidence#OR2.3.3.5_HelpDesk Form*	Peers recommend that ODE provide a basic troubleshooting guide. Evidence includes numerous references to resources that users can contact for technical support (e.g., Regional ESD Partners, OSAS Help Desk, ODE Policy Contacts). It would be helpful for the state to provide a list or chart of some type that puts in one place which contact to call for the different categories of issues that might arise. Otherwise, when TAs or STCs are in the thick of testing and an issue arises that needs immediate attention, it is often confusing to know whether to contact the district assessment office, the state assessment office, or the vendor tech support line.

Section 2.3 Summary Statement

___ No additional evidence is required or

_X__ The following additional evidence is needed/provide brief rationale:

- Evidence that the State has established procedures to ensure that general and special education teachers, paraprofessionals, teachers of ELs, specialized instructional support personnel, and other appropriate staff receive necessary training to administer assessments and know how to administer assessments, including, as necessary, alternate assessments, and know how to make use of appropriate accommodations during assessments for all students with disabilities and English learners.

Critical Element 2.4 – Monitoring Test Administration

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or 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. Monitoring of test administration should be demonstrated for all assessments in the State system: the general academic assessments and the AA-AAAS.	Reviewed by Department Staff Only	Department staff determined that the State's evidence is not sufficient for this critical element. The State must provide evidence that it adequately monitors the administration of its State assessments to ensure that standardized test administration procedures are implemented with fidelity across districts and schools.

Section 2.4 Summary Statement

___ No additional evidence is required or

_X__ The following additional evidence is needed/provide brief rationale:

- Evidence that the State adequately monitors the test administration of its State assessments to ensure that standardized test administration procedures are implemented with fidelity across districts and schools (e.g., evidence of scheduled visits with district or schools, or redacted letters sent to districts or schools following the in-person monitoring visits).

Critical Element 2.5 – Test Security

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or 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: • 2.5.1: Prevention of any assessment irregularities, including maintaining the security of test materials (both during test development and at time of test administration), 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; • 2.5.2: Detection of test irregularities;	***2.5.1*** *• _Please see Common Peer Review Submission for Science Supplemental Evidence and Notes* *• _Evidence#OR1.3.1_2018_19_TAM* *• _Evidence#OR2.3.1.1_DTC Training Materials Webscreen and Training Modules* *• _Evidence#OR2.4.2_2018_19 Redacted Science Test Impropriety Log* *• _Evidence#OR2.4.3_RedactedSampleLetterofFinalDetermination* *• _Evidence#OR2.1.1.1_V1_OSAS Science Assessment Tech Report 2018-2019* *• _Evidence#OR2.5.1.1_TestImpropIrregRolesResp* • *Evidence#OR1.5.4.2_V3_OSAS Science Assessment Tech Report_v2NG*, pdf p. 22 (added by peers) **2.5.2** *Please see Common Peer Review Submission for Science Supplemental Evidence and Notes* *• _Evidence#OR1.3.1_2018_19_TAM* *• _Evidence#OR2.4.1_TestImproprietyForm* *(Did not find evidence of detection of irregularities in the above sources given by OR)* (*Evidence#OR2.1.1.1_V1_OSAS Science Assessment Tech Report 2018-2019*, pdf p 68). added by peers	**CE 2.5.1:** *Peers found evidence that the State has procedures in place for the prevention of assessment irregularities, including maintaining the security of test materials (both during test development and at time of test administration), 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.* **CE 2.5.2**: *Peers did **not** find sufficient evidence that 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 detection of test irregularities.* Evidence that the state could provide to meet this CE includes evidence of data forensics designed to identify testing irregularities or cheating at the student, class, or teacher level. While peers note the item analysis report (*Evidence#OR2.1.1.1_V1_OSAS Science Assessment Tech Report 2018-2019*, pdf p 68) and the web-based irregularity detection by Caveon

		(*Evidence#OR2.1.1.1_V1_OSAS Science Assessment Tech Report 2018-2019, Appendix F)*, this evidence does not provide sufficient information on detection of test irregularities at the levels described above.
• 2.5.3: Remediation following any test security incidents involving any of the State's assessments;	**CE 2.5.3 & C 2.5.4** *Please see Common Peer Review Submission for Science Supplemental Evidence and Notes* **(2.5.3)** • *_Evidence#OR1.3.1_2018_19_TAM*	**CE 2.5.3**: *Peers found sufficient evidence of procedures for remediation following any test security incidents involving any of the State's assessments* Peers note that it is unclear what remediation has taken place or will take place at the state level when patterns of irregularities occur across the state.
• 2.5.4: Investigation of alleged or factual test irregularities.		**CE 2.5.4:** *Peers did **not** find sufficient evidence of procedures for investigation of alleged or factual test irregularities.* While evidence is provided in a general form regarding procedures for test irregularities, peers note that there is no guidance for DTCs on specific steps to take/how to conduct the investigation. It looks like the ODE relies on the district to do so, but there is no guidance provided on whom to interview, format for reporting or how much detail to provide to state, etc. Given that licensure action is listed as one of the potential consequences of testing irregularities, it seems likely that the State Board attorneys need certain types of information for these cases, and peers did not find appropriate documentation around this.
• 2.5.5: Application of test security procedures to all assessments in the State system: the general academic assessments and the AA-AAAS.	**CE 2.5.5** • *_Please see common submission for supplemental evidence and notes.* • *_Evidence#OR1.3.1_2018_19_TAM* • *_Evidence#OR2.4.1_TestImproprietyForm* • *Evidence#OR2.4.2_2018_19 Redacted Science Test Impropriety Log* • *Evidence#OR2.4.3_RedactedSampleLetterofFinalDetermination* • *Evidence#OR1.4.2_2018-19_OAM_Final (added by peers)*	**CE 2.5.5:** *Peers found evidence of application of test security procedures to all assessments in the State system: the general academic assessments and the AA-AAAS.* *Evidence#OR1.4.2_2018-19_OAM_Final,* p. 47 notes that "appropriate test security" needs to be followed when a human reader is involved in test administration. Peers note that this guidance could be more specific.

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Section 2.5 Summary Statement

___ No additional evidence is required or

_X__ The following additional evidence is needed/provide brief rationale:

- Evidence that 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 detection of test irregularities.
- Evidence of specific procedures for investigation of alleged or factual test irregularities.

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 2.6 – Systems for Protecting Data Integrity and Privacy

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or 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:	**CE 2.6**	
• 2.6.1: To protect the integrity of its test-related data in test administration, scoring, storage and use of results;	***2.6.1*** • *_Please see common submission for supplemental evidence and notes.* • *_Evidence#OR1.3.1_2018_19_TAM* • *_Evidence#OR2.6.1.1_ORS336_184* • *_Evidence#OR2.6.1.2_AIRContract*	**CE 2.6.1**: *Peers found evidence that the State has procedures in place to protect the integrity of its test-related data in test administration, scoring, storage and use of results.*
• 2.6.2: To secure student-level assessment data and protect student privacy and confidentiality, including guidelines for districts and schools;	***2.6.2*** • *_Please see common submission for supplemental evidence and notes.* • *Evidence#OR1.3.1_2018_19_TAM* • *_Evidence#OR2.6.1.1_ORS 336_184* • *_Evidence#OR2.6.2.1_HB2715* • *_Evidence#OR2.6.2.2_ExecutiveNumberedMemo005-2015-16* • *_Evidence#OR2.6.2.3_ODEPolicy_581-101* • *_Evidence#OR2.6.2.4_ODEPolicy_581-116*	**CE 2.6.2:** *Peers found evidence that the State has procedures to secure student-level assessment data and protect student privacy and confidentiality, including guidelines for districts and schools.*
• 2.6.3: 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.	***2.6.3*** • *_Please see common submission for supplemental evidence and notes.* • *_Evidence#OR2.6.2.4_ODEPolicy_581-101 (This is not correctly numbered)* • *_Evidence#OR2.6.2.1_ORS 336_184* • *_Evidence#OR2.6.2.2_HB2715* • *_Evidence#OR2.6.2.3_ExecutiveNumberedMemo005-2015-16 (not correctly numbered)* (*Evidence#OR2.6.2.3_ODEPolicy_581-101*) – this is the correct numbering for this document based on the list of evidence to peers	**CE 2.6.3**: *Peers found evidence that the State has procedures 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#OR2.6.1.1_ORS 336_184 describes the Oregon Student Information Protection Act, which includes student records and explicitly identifies test results.

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

		"For groups where 5 or fewer students can be isolated, there is potential for identification of an individual student. In these cases, the student counts or percents should be masked. In the case of percents as in the Assessment Results where 100 percent of a student population fall into a result category the data will be masked." (*Evidence#OR2.6.2.3_ODEPolicy_581-101,* p.4)

Section 2.6 Summary Statement

_X__ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:

- [list additional evidence needed w/brief rationale]

SECTION 3: TECHNICAL QUALITY – VALIDITY

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

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State has documented adequate overall validity evidence for its assessments consistent with nationally recognized professional and technical testing standards. The State's validity evidence includes evidence that: **<u>The State's academic assessments</u>** measure the knowledge and skills specified in the State's academic content standards, including: • 3.1.1 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), balance of content, and cognitive complexity;	**CE 3.1** *Please see common submission for supplemental evidence and notes.* ***3.1.1*** • *_Evidence#OR2.1.2.6_V4 OSAS Science Tech Report 2018-2019* • *Evidence#OR3.1.1.1_CategoryEngagementDefinitions* • *_Evidence#OR3.1.1.2_Oregon Supplemental Science Peer Review Evidence* .(*Evidence#OR1.5.4.1_V2_OSAS Science Assessment Technical Report 2018-2019,* pdf pg. 24) (added by peers) • *A_Oregon2018_19SciencePeerReviewIndex_VF, CE 4.4, pdf pg. 65, pg. 62* (added by peers)	**CE 3.1.1:** *Peers did **not** find sufficient evidence of 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), balance of content, and cognitive complexity.* Peers note that to fully meet state expectations, at least 25 items would need to be added that address otherwise unrepresented PEs. (*Evidence#OR2.1.2.6_V4 OSAS Science Tech Report 2018-2019,* pdf pg. 219). Plans to develop additional items to address the underrepresented PEs are described in *Evidence#OR3.1.1.2_Oregon Supplemental Science Peer Review Evidence.* With regard to cognitive complexity, *OR2.1.2.6_V4 OSAS Science Tech Report 2018-2019*, pdf p. 217 states, "Category 4 expectations, which are complex tasks that require extended time (such as the "sustained investigations" expected by the Framework) are not expected to be appropriately or authentically assessed in an on-demand context." However, there were minimal

		Category 4 items available in the Shared Item Bank, but since Oregon chose to eliminate all constructed-response items from their state assessment (*Oregon State Index, CE 4.4*), these opportunities for Level 4 items were lost. In addition, other large-scale on-demand assessments have been developed with the inclusion of extended performance tasks (e.g., SBA) that assess Category 4 level complexity. Therefore, the state is encouraged to revisit their decision to eliminate constructed-response items from their assessment. Otherwise the full extent of cognitive complexity is not measured.
• 3.1.2 Documentation that the assessments address the depth and breadth of the content standards;	**3.1.2** • *_Evidence#OR2.1.2.6_V4_OSAS Science Tech Report 2018-2019* • *A_Oregon2018_19SciencePeerReviewIndex_VF, CE 4.4, pdf pg. 65, pg. 62* (added by peers)	**CE 3.1.2:** *Peers did **not** find sufficient evidence of documentation that the assessments address the depth and breadth of the content standards.* If the assessment scores are providing information about student strengths and weaknesses only in terms of discipline levels (i.e., science content), then the assessment is not fully measuring the depth and breadth of the NGSS, since the standards are three-dimensional, with DCI (content) being only one of three dimensions. The scores should also provide evidence in terms of students' abilities in the use of Science and Engineering Practices (SEPs) and Cross-Cutting Concepts (CCCs), in order to truly support students' science learning in the context of the NGSS. In addition, by Oregon having removed all constructed-response items in the bank from their state assessment, (*Oregon State Index, CE 4.4*), the assessment is less likely to be adequately measuring students' communication of science, a key piece in at least three of the eight SEPs of the NGSS.
• 3.1.3 If the State has adopted alternate academic achievement standards and administers alternate assessments aligned with those standards, the assessments show adequate alignment to the State's academic content standards for the grade in which the student is enrolled 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.	***3.1.3*** *Not applicable*	**CE 3.1.3:** This aspect of the CE is not applicable to this submission.

Section 3.1 Summary Statement		
___ No additional evidence is required or X___ The following additional evidence is needed/provide brief rationale: • Evidence of documentation that the assessments address the depth and breadth of the content standards. • Evidence of 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), balance of content, and cognitive complexity.		

Critical Element 3.2 – Validity Based on Cognitive Processes

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or 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.	**CE 3.2** *Please see common submission for supplemental evidence and notes.* • *_Evidence#OR2.1.2.6_V4_OSAS Science Tech Report 2018-2019* • *Evidence#OR3.1.1.1_CategoryEngagementDefinitions*	**CE 3.2:** *Peers found evidence that 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.* Comments made above in 3.1.1 point out the need for the state and/or consortium to develop items that tap cognitive Complexity Level 4 to ensure that the assessment taps all the intended cognitive processes as expected in the NGSS.

Section 3.2 Summary Statement

_X__ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:

- [list additional evidence needed w/brief rationale]

Critical Element 3.3 – Validity Based on Internal Structure

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or 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**.	**CE 3.3** *Please see common submission for supplemental evidence and notes.* • *_Evidence#OR2.1.1.1_V1_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR2.1.2.6_V4_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR3.1.1.2_Oregon Supplemental Science Peer Review Evidence*	**CE 3.3:** *Peers did **not** find that 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**.* Regarding evidence #3.1.1.2, it is not clear why OR conducted a Principal Components factor analysis in a somewhat exploratory way, rather than directly doing a Confirmatory factor analysis (CFA), to support a hypothesis of a single factor with X number of subdomains. Slides 5, 6, 7. The state evidence describes reporting scores in terms of both science disciplines and in terms of DCIs. However, the key structure of the NGSS is three-dimensional using DCIs, SEPs, and CCCs. No validity evidence is found documenting that the scoring and reporting structure of its assessment is consistent with the three-dimensional structure of the State's standards. Peers recommend that the State works with the consortium on this aspect (since it is an integral component of test form and item development). *Evidence#OR2.1.1.1_V1_OSAS Science Assessment Tech Report 2018-2019,* Section 6.7. Please see consortium submission feedback for concerns about the subdomains used for this assessment. Peers believe that, in order for the assessment to have validity in relation to the academic content standards, the subdomains need to be defined in terms of the multi-dimensional nature of the NGSS (DCI, SEP, CCC). Evidence that could help meet this requirement includes a rationale to support the use of content-specific subdomains (rather than the three dimensions of the NGSS) and a CFA

		investigating whether the items sufficiently support reporting at the SEP and CCC level.

Section 3.3 Summary Statement

___ No additional evidence is required or

_X__ The following additional evidence is needed/provide brief rationale:

- Evidence documenting that the scoring and reporting structure of the assessment is consistent with the three-dimensional nature of the State's academic content standards (NGSS).

Critical Element 3.4 – Validity Based on Relations to Other Variables

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State has documented adequate validity evidence that the State's assessment scores are related as expected with other variables.	**CE 3.4** • *_Please see common submission for supplemental evidence and notes.* • *_Evidence#OR2.1.2.6_V4_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR3.1.1.2_Oregon Supplemental Science Peer Review Evidence*	**CE 3.4***: Peers found evidence that the State has documented adequate validity evidence that the State's assessment scores are related as expected with other variables.* Panel expects that further research will be done as more data are available (peers noted that there was planned research that was disrupted due to the pandemic). Provided correlations with other OR assessments. Interruption in research due to pandemic.

Section 3.4 Summary Statement

__X_ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:

- [list additional evidence needed w/brief rationale]

SECTION 4: TECHNICAL QUALITY – OTHER

Critical Element 4.1 – Reliability

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or 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 consistent with nationally recognized professional and technical testing standards. If the State's assessments are implemented in multiple States, measures of reliability for the assessment overall and each student group consistent with nationally recognized professional and technical testing standards, including:		
• 4.1.1: Test reliability of the State's assessments estimated for its student population;	***4.1.1*** • *_Please see common submission for supplemental evidence and notes.* • *_Evidence#OR2.1.2.6_V4_OSAS Science Assessment Tech Report 2018-2019*	**CE 4.1.1:** *Peers did **not** find sufficient evidence that the state has adequately documented test reliability of the State's assessments estimated for its student population;* Without numbers for sample size of students (table A-1, pg. 36 of 2.1.2.6), reliability is unclear. Peers request that the state include the number of students in subgroups so that peers can understand whether or not reliability is acceptable (consistent with nationally recognized assessment standards) and, if reliability is not acceptable, a plan for improvement of reliability is requested
• 4.1.2: Overall and conditional standard error of measurement of the State's assessments, including any domain or component sub-tests, as applicable;	***4.1.2*** • *_Evidence#OR2.1.2.6_V4_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR3.1.1.2_Oregon Supplemental Science Peer Review Evidence*	**CE 4.1.2:** *Peers found evidence that the state has adequately documented overall and conditional standard error of measurement of the State's assessments, including any domain or component sub-tests, as applicable* Peers recommend that additional item development include items that specifically target high- and low-achieving students.

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
• 4.1.3: Consistency and accuracy of estimates in categorical classification decisions for the cut scores, achievement levels or proficiency levels based on the assessment results;	***4.1.3*** • *_Evidence#OR2.1.2.6_V4_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR3.1.1.2_Oregon Supplemental Science Peer Review Evidence*	**CE 4.1.3:** *Peers found evidence that the state has adequately documented consistency and accuracy of estimates in categorical classification decisions for the cut scores, achievement levels or proficiency levels based on the assessment results.*
• 4.1.4: For computer-adaptive tests, evidence that the assessments produce test forms with adequately precise estimates of a **student's academic achievement**.	***4.1.4*** • *_Evidence#OR2.1.2.6_V4_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR3.1.1.2_Oregon Supplemental Science Peer Review Evidence*	**CE 4.1.4:** *Peers found evidence that the state has adequately documented for computer-adaptive tests, evidence that the assessments produce test forms with adequately precise estimates of a **student's academic achievement**.*

Section 4.1 Summary Statement

___ No additional evidence is required or

_X__ The following additional evidence is needed/provide brief rationale:

- Evidence that includes the number of students in subgroups in order to demonstrate that reliability is at an acceptable level for subgroups (consistent with nationally recognized assessment standards) and, if reliability is not acceptable, a plan for improvement of reliability.

Critical Element 4.2 – Fairness and Accessibility

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
For all State academic assessments, assessments should be developed, to the extent practicable, using the principles of universal design for learning (UDL) (see definition[1]). **For academic content assessments**, the State has taken reasonable and appropriate steps to ensure that its assessments are accessible to all students and fair across student groups in their design, development and analysis.	**CE 4.2** • *_Please see Common Peer Review Submission for Science Supplemental Evidence and Notes.* • *_Evidence#OR1.5.4.1_V2_OSAS Science Assessment Technical Report 2018-2019* • *_Evidence#OR4.2.1_SAIC Item Specifications Guidelines* • *_Evidence#OR4.2.2_Appendix A. Classical Statistics for Science Items* • *_Evidence#OR2.1.1.1_V1_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR2.1.2.6_V4_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR4.2.3_V5_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR3.1.1.2_Oregon Supplemental Science Peer Review Evidence* • *A_Oregon2018_19SciencePeerReviewIndex_VF, CE 4.4, pdf pg. 65, pg. 62* (added by peers)	**CE 4.2:** *Peers found sufficient evidence that the State has taken reasonable and appropriate steps to ensure that its assessments are accessible to all students and fair across student groups in their design, development and analysis.* As noted in the comment in CE 2.3.1 in this state report, Oregon's instructions to TAs simply "encourage" students be given the practice test, but do not explicitly require it. It is essential for fairness to all students that provisions are in place for every student to take the practice test every year they are tested on this assessment, and for a system to be in place to document that the practice test was in fact taken by every student prior to testing. • Comments on the Common submission review pointed out a strength in the item development process as utilizing a wide variety of item response types ranging from selected-response to various technology-enhanced items, as well as constructed response items, thereby reflecting UDL by allowing students a variety of ways to show what they know and can do. Oregon has eliminated from their state assessment all constructed-response items from the bank. As a result, because some students are better able to show what they know and can do through constructed responses, this restriction in the Oregon assessment lowers the likelihood that the assessment is accessible to all students. Peers recommend the inclusion of extended response items in the assessment. (*A_Oregon2018_19SciencePeerReviewIndex_VF, CE 4.4, pdf pg. 65, pg. 62)*

[1] see page 28 of "*A State's Guide to the U.S. Department of Education's Assessment Peer Review Process"*, September 24, 2018 available at: www.ed.gov/admins/lead/account/saa.html

Section 4.2 Summary Statement		
_X__ No additional evidence is required or ___ The following additional evidence is needed/provide brief rationale: • [list additional evidence needed w/brief rationale]		

Critical Element 4.3 – Full Performance Continuum

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State has ensured that each assessment provides an adequately precise estimate of student performance across the full performance continuum for **academic assessments**, including performance for high- and low-achieving students.	**CE 4.3** • *_Please see common submission for supplemental evidence and notes.* • *_Evidence#OR4.3.1_Appendix B_Estimated Item and Group Parameters* • *_Evidence#OR2.1.1.1_V1_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR2.1.2.6_V4_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR3.1.1.2_Oregon Supplemental Science Peer Review Evidence*	**CE 4.3:** *Peers found evidence that the State has ensured that each assessment provides an adequately precise estimate of student performance across the full performance continuum for academic assessments, including performance for high- and low-achieving students.* The percent of students scoring in level 4 is not consistent with other Oregon state assessments, particularly grade 8, which shows 2% of students scoring at level 4, as opposed to roughly 20% of grade 8 Oregon students scoring at level 4 on the mathematics assessment. The omission of extended response items from the science assessment may be linked to the dearth of items that are able to differentiate level 4 performance from level 3 performance. Peers question whether the distribution of scores by level is acceptable and recommend the development of additional items at upper levels, to distinguish performance between levels 3 and 4. Peers recommend that Oregon consider evaluating the proportion of Oregon students scoring at the highest level (level 4) compared to those states making use of the full range of item types from the consortium.

Section 4.3 Summary Statement

_X__ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:

- [list additional evidence needed w/brief rationale]

Critical Element 4.4 – Scoring

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State has established and documented standardized scoring procedures and protocols for its assessments that are designed to produce reliable and meaningful results, facilitate valid score interpretations, and report assessment results in terms of the State's **academic achievement standards**.	**CE 4.4** Please see *Common Peer Review Submission for Science Supplemental Evidence and Notes*. • *_Evidence#OR2.1.1.1_V1_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR1.5.4.1_ V2_OSAS Science Assessment Tech Report* • *_Evidence#OR3.1.1_Oregon Supplemental Science Peer Review Evidence* • *_Evidence#OR4.4.1_MMLE Validation Study* • *_Evidence#OR1.3.1_2018_19_TAM*	**CE 4.4:** *Peers found evidence that the State has established and documented standardized scoring procedures and protocols for its assessments that are designed to produce reliable and meaningful results, facilitate valid score interpretations, and report assessment results in terms of the State's* ***academic achievement standards.*** Peers found that the assessment has technically met the CE, but conceptually, the interpretation of the scores will be limited by the following: - One of the stated purposes of the OR science assessment is to "support instruction and student learning by measuring growth in student achievement…. Test scores can be employed to evaluate students' learning progress and to help teachers to improve their instruction, which in turn has a positive effect on students' learning over time." (*CS 004 - V4_Shared Technical Report,* pdf p. 9). It is not clear how the scores will be used to measure student growth in science achievement over time. Peers found an apparent misalignment between scoring and the intended use of results – vertically articulated scores/standards are not evident. It is also difficult to measure **growth** with assessments administered every three years. - In addition, as noted in the Common submission, scores for this science assessment are not being reported in terms of the three dimensions of the NGSS, but rather in terms of simply the science content (DCI) dimension.

Section 4.4 Summary Statement

_X__ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:
- [list additional evidence needed w/brief rationale]

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

STATE ASSESSMENT PEER REVIEW NOTES FOR Oregon - Science

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 4.5 – Multiple Assessment Forms

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
If the State administers multiple forms of **academic assessments** 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.	**CE 4.5** • *_Please see common submission for supplemental evidence and notes.* • *_Evidence#OR4.5.1_LOFT and CAT Designs* • *_Evidence#OR2.1.1.1_V1_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR1.5.4.1_V2_OSAS Science Assessment Technical Report 2018-2019* • *Evidence #OR2.3.1.1_DTC Training Materials Webscreen and Training Modules*, pg. 25 (added by peers)	**CE 4.5:** *Peers did **not** find sufficient evidence that 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.* Peers recommend that the state provide evidence that the Braille form is comparable with the general education assessment, given that the blueprint was not met for the Braille forms. Peers found other requirements of this CE are met by the submission. The simulation that was conducted did not examine the fixed linear Braille forms. Oregon intends to use the LOFT test design in 2020-21 to select stand-alone and cluster items for students who need the Braille version. The state will need to do further work to validate that students' scores on the Braille version are comparable to those of students on other versions.

Section 4.5 Summary Statement

___ No additional evidence is required or

_X__ The following additional evidence is needed/provide brief rationale:

- Evidence that the Braille blueprint is comparable to the general education assessment blueprint.

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 4.6 – Multiple Versions of an Assessment

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
If the State administers any of its assessments in multiple versions within a subject area (e.g., online versus paper-based delivery; **or a native language version of the academic content assessment**), grade level, or school year, the State: • 4.6.1: Followed a design and development process to support comparable interpretations of results for students tested across the versions of the assessments;	**CE 4.6** ***4.6.1*** • *_Please see common submission for supplemental evidence and notes.* • *_Evidence#OR2.1.1.1_V1_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR1.5.4.1_V2_OSAS Science Assessment Technical Report 2018-2019* • *_Evidence#OR4.2.3_V5_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR2.1.2.6_V4_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR4.6.1.1_Smarter Balanced Translation Framework*	**CE 4.6.1**: *Peers found evidence that the State followed a design and development process to support comparable interpretations of results for students tested across the versions of the assessments.*
• 4.6.2: Documented adequate evidence of comparability of the meaning and interpretations of the assessment results.	***4.6.2*** • *_Evidence#OR3.1.1.2_Oregon Supplemental Science Peer Review Evidence* • *_Evidence#OR2.1.1.1_V1_OSAS Science Assessment Tech Report*	**CE 4.6.2**: *Peers did **not** find evidence that the State documented adequate evidence of comparability of the meaning and interpretations of the assessment results*. Peers note that Oregon used the SBAC framework but do not see evidence that this translation framework results in appropriate comparability among versions. Peers request results of the planned comparability studies using student performance across the versions of the assessment. While the state followed a design and development process to show content comparability across the major versions of the assessment, no evidence is found of actual comparability studies using student results to validate

		comparability of the meaning and interpretations of the assessment results across different versions of the test. In addition, evidence is needed regarding the administration of the assessment using different types of devices (e.g., desktop computers, laptops, tablets). Specific evidence should document that test-administration hardware and software (e.g., screen resolution, interface, input devices) are standardized across unaccommodated administrations; or (a) research reports (quantitative or qualitative) that show that variations resulting from different types of delivery devices do not alter the interpretations of results; or (b) an appropriate comparability study.

Section 4.6 Summary Statement

___ No additional evidence is required or

_x__ The following additional evidence is needed/provide brief rationale:

- Comparability studies across different versions of the test based on student performance.

Critical Element 4.7 – Technical Analysis and Ongoing Maintenance

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State: • 4.7.1: Has a system for monitoring, 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), and • 4.7.2: Evidence of adequate technical quality is made public, including on the State's website.	**CE 4.7** • *_Please see common submission for supplemental evidence and notes.* • *_Evidence#OR2.1.1.1_V1_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR1.5.4.1_V2_OSAS Science Assessment Technical Report 2018-2019* • *_Evidence#OR1.5.4.2_V3_OSAS Science Tech Report* • *_Evidence#OR2.1.2.6_V4_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR4.2.3_V5_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR4.7.1_V6_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR4.7.2_ Assessment Technical Manuals* • *_Evidence#OR3.1.1.2_Oregon Supplemental Science Peer Review Evidence* • *A_Oregon2018_19SciencePeerReviewIndex_VF, CE 4.4, pdf pg. 65, pg. 62* (added by peers)	**CE 4.7.1**: *Peers did **not** find sufficient evidence that the State has a system for monitoring, maintaining, and improving, as needed, the quality of its assessment system, including clear and technically sound criteria for the analyses of the state science assessment.* Peers could not locate information describing the composition of the state Technical Advisory Committee or how this entity is used for monitoring, maintaining, and improving the quality of the state science assessment system. Evidence is provided of various planned studies and analyses designed to monitor the quality of the science assessment. However, it is not clear how the state will be using the information gained during monitoring to improve the assessment for the future (e.g., who will do the improvements, the timeline). **CE 4.7.2:** *Peers did **not f**ind that evidence of adequate technical quality is made public, including on the State's website.* Peers note that, in Oregon's submission, it is stated that the technical reports are still being prepared for public presentation. Therefore, this evidence of adequate technical quality is not currently made public. There is currently a lack of clarity in the technical manual about what is true for the overall consortium, in terms of development and design, and what is true for the State-specific assessment. For example, peers recommend that the State include information in the technical reports about the actual item types/interaction types included in the assessment (currently item types not included on the assessment, such as extended response, are listed in the

		technical report *Evidence#OR1.5.4.1_V2_OSAS Science Assessment Tech Report*, pdf p. 24). The 8th paragraph of the State's Index, under CE 4.4, mentions that there are no constructed-response items on the test. However, this is not made public, and, conversely, evidence that leads the reader to believe the test consists of constructed-response items is populated throughout the documentation.

Section 4.7 Summary Statement

___ No additional evidence is required or

_X_ The following additional evidence is needed/provide brief rationale:

- Evidence of adequate technical quality is made public, including on the State's website, and information in the technical manuals clarifies what is true for the Consortium compared to what is true for the state assessment.
- Evidence that the State has a system for monitoring, maintaining, and improving, as needed, the quality of its assessment system, including clear and technically sound criteria for the analyses of the state science assessments (e.g. composition and use of the technical advisory committee, how planned studies will be used to monitor and improve the quality of the assessments).

SECTION 5: INCLUSION OF ALL STUDENTS

Critical Element 5.1 – Procedures for Including Students with Disabilities

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
5.1.1: 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. Decisions about how to assess students with disabilities must be made by a student's IEP Team under IDEA, the placement team under Section 504, or the individual or team designated by a district to make that decision under Title II of the ADA, as applicable, based on each student's individual abilities and needs. 5.1.2: If a State adopts alternate academic achievement standards for students with the most significant cognitive disabilities and administers an alternate assessment aligned with those standards under ESEA section 1111(b)(1)(E) and (b)(2)(D), respectively, the State must: • Establish guidelines for determining whether to assess a student with an AA-AAAS, including: o A State definition of "students with the most significant cognitive disabilities" that addresses factors related to cognitive functioning and adaptive behavior; • Provide information for IEP Teams to inform decisions about student assessments that:	**CE 5.1** • *_Evidence#OR1.3.1_2018_19_TAM* • *_Evidence#OR1.4.2_2018_19_OAM_Final* • *_Evidence#OR5.1.1_OregonIEP*	CE 5.1.1: *Peers found evidence that 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. Decisions about how to assess students with disabilities must be made by a student's IEP Team under IDEA, the placement team under Section 504, or the individual or team designated by a district to make that decision under Title II of the ADA, as applicable, based on each student's individual abilities and needs.* *CE 5.1.2: Does not apply.*

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
○ Provides a clear explanation of the differences between assessments aligned with grade-level academic achievement standards and those aligned with alternate academic achievement standards, including any effects of State and local policies on a student's education resulting from taking an AA-AAAS, such as how participation in such assessments may delay or otherwise affect the student from completing the requirements for a regular high school diploma; • Ensure that parents of students assessed with an AA-AAAS are informed that their child's achievement will be measured based on alternate academic achievement standards; • Not preclude a student with the most significant cognitive disabilities who takes an AA-AAAS from attempting to complete the requirements for a regular high school diploma; and • Promote, consistent with requirements under the IDEA, the involvement and progress of students with the most significant cognitive disabilities in the general education curriculum that is based on the State's academic content standards		

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
for the grade in which the student is enrolled; and • Develop, disseminate information on, and promote the use of appropriate accommodations to ensure that a student with the most significant cognitive disabilities who does not take an AA-AAAS participates in academic instruction and assessments for the grade in which the student is enrolled. • The State has in place and monitors implementation of guidelines for IEP teams to apply in determining, on a case-by-case basis, which students with the most significant cognitive disabilities will be assessed based on alternate academic achievement standards, if applicable. Such guidelines must be developed in accordance with 34 CFR § 200.6(d).[2]		

Section 5.1 Summary Statement

__X_ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:
- [list additional evidence needed w/brief rationale]

[2] See the full regulation at 34 CFR § 200.6(d) (online at https://www.ecfr.gov/cgi-bin/text-idx?SID=07e168e9e7a6c5931b4549cc15547ee9&mc=true&node=se34.1.200_16&rgn=div8)
Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 5.2 – Procedures for Including English Learners in Academic Content Assessments

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State has in place procedures to ensure the inclusion of all ELs in public elementary and secondary schools in the State's academic content assessments and clearly communicates this information to districts, schools, teachers, and parents, including, at a minimum: • 5.2.1 Procedures for determining whether an EL should be assessed with a linguistic accommodation(s); • 5.2.2 Information on accessibility tools and features available to all students and assessment accommodations available for ELs; • 5.2.3 Assistance regarding selection of appropriate linguistic accommodations for ELs, including to the extent practicable, assessments in the language most likely to yield accurate and reliable information on what those students know and can do to determine the students' mastery of skills in academic content areas until the students have achieved English language proficiency.	**CE 5.2** • _Evidence *#OR1.4.2_2018_19_OAM_Final*	**CE 5.2.1:** *Peers found evidence that the State has in place procedures for determining whether an EL should be assessed with a linguistic accommodation(s).* **CE 5.2.2:** *Peers found evidence that the State has in place information on accessibility tools and features available to all students and assessment accommodations available for ELs.* **CE 5.2.3:** *Peers found evidence that the State has in place assistance regarding selection of appropriate linguistic accommodations for ELs, including to the extent practicable, assessments in the language most likely to yield accurate and reliable information on what those students know and can do to determine the students' mastery of skills in academic content areas until the students have achieved English language proficiency.* Peers recommend providing additional supports for students who speak languages other than Spanish. Peers recommend that the state consider the use of word-to-word dictionaries for EL students (see note in CE 5.3 regarding the need for this accommodation).
Section 5.2 Summary Statement		
_X__ No additional evidence is required or		

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

STATE ASSESSMENT PEER REVIEW NOTES FOR Oregon - Science

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
___ The following additional evidence is needed/provide brief rationale: • [list additional evidence needed w/brief rationale]		

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 5.3 – Accommodations

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State makes available appropriate accommodations and ensures that its assessments are accessible to students with disabilities and ELs, including ELs with disabilities. Specifically, the State: • 5.3.1: Ensures that appropriate accommodations, such as, interoperability with, and ability to use, assistive technology, are available to measure the **academic achievement** of students with disabilities. • 5.3.2: Ensures that appropriate accommodations are available for ELs;	**CE 5.3** • *_Please see common submission for supplemental evidence and notes.* • *_Evidence#OR1.3.1_2018_19_TAM* • *_Evidence#OR1.4.2_2018_19_OAM_Final* • *_Evidence#OR2.1.1.1_V1_OSAS Science Assessment Tech Report 2018-2019*	**CE 5.3.1:** *Peers found evidence that the State ensures that appropriate accommodations, such as, interoperability with, and ability to use, assistive technology, are available to measure the **academic achievement** of students with disabilities.* **CE 5.3.2:** *Peers did **not** find sufficient evidence that the State ensures that appropriate accommodations are available for ELs.* The Common submission indicates that CAI allows a bilingual/dual language word-to-word dictionary as a non-embedded support for students whose primary language is not English and who use dual language supports in the classroom. This support is not listed as being allowed on the Oregon version of the assessment. It is unclear why such a support has not been included in the Oregon assessment. In addition, its exclusion causes concern in terms of comparability statistics for ELs generated by CAI on the overall item bank and those for ELs generated by studies conducted by Oregon on their version of the assessment. It is recommended that either (a) Oregon adds the use of a bilingual/dual language word-to-word dictionary to their allowable supports, or (b) a study be conducted by the state to validate that the results on the assessment for ELs are equivalent with or without such a support.

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
• 5.3.3: Has determined that the accommodations it provides (1) are appropriate and effective for meeting the individual student's need(s) to participate in the assessments, (2) do not alter the construct being assessed, and (3) allow meaningful interpretations of results and comparison of scores for students who need and receive accommodations and students who do not need and do not receive accommodations;		**CE 5.3.3:** *Peers did **not** find sufficient evidence that the State has determined that the accommodations it provides (1) are appropriate and effective for meeting the individual student's need(s) to participate in the assessments, (2) do not alter the construct being assessed, and (3) allow meaningful interpretations of results and comparison of scores for students who need and receive accommodations and students who do not need and do not receive accommodations* In the review of the Common submission, it was noted that: o For Spanish-speaking ELs, no evidence is provided regarding the review of the Spanish translations by native Spanish speakers. o For visually-impaired students, the Braille version suffered from a number of formatting and technical issues that bring into question whether the results for these students are comparable to results for students not taking the Braille version. No evidence is found that the state addressed these issues for their state version of the assessment. The lack of the use of a bilingual/dual language word-to-word dictionary for ELs may be altering the construct being assessed, since the goal is to assess students' science achievement, not their ability to understand English vocabulary that is unrelated to the science concept being tested. It appears that there is no paper/pencil version of the science assessment for Oregon. This may be due to the fact that Oregon has eliminated all the open-ended items that are in the shared item bank from their version of the assessment. It is presumed that any student who would have been given a paper/pencil version will receive the scribe accommodation (human assist to enter student's

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
• 5.3.4: Has a process to individually review and allow exceptional requests for a small number of students who require accommodations beyond those routinely allowed. • 5.3.5: Ensures that accommodations for all required assessments do not deny students with disabilities or ELs the opportunity to participate in the assessment and any benefits from participation in the assessment.		responses). While this is acceptable, it can sometimes be more difficult or awkward for a student to use a scribe rather than putting their responses on paper, and it opens the process to more potential for loss of security or validity even with trained TAs. It is not clear if students who use the print-on-demand accommodation might be able to mark their answers on the print document, which the TA would then enter into the online system. Peers did not find evidence of comparability studies on performance of students who use accommodations vs. those who do not. **CE 5.3.4:** *Peers found evidence that the State has a process to individually review and allow exceptional requests for a small number of students who require accommodations beyond those routinely allowed.* **CE 5.3.5:** *Peers found evidence that the State ensures that accommodations for all required assessments do not deny students with disabilities or ELs the opportunity to participate in the assessment and any benefits from participation in the assessment.*

Section 5.3 Summary Statement

___ No additional evidence is required or

_X__ The following additional evidence is needed/provide brief rationale:

- Evidence that either (a) Oregon adds the use of a bilingual/dual language word-to-word dictionary to their allowable supports, or (b) a study be conducted by the state to validate that the results on the assessment for ELs is equivalent with or without such a support.
- Evidence that the State has determined that the accommodations it provides (1) are appropriate and effective for meeting the individual student's need(s) to participate in the assessments, (2) do not alter the construct being assessed, and (3) allow meaningful interpretations of results and comparison of scores for students who need and receive accommodations and students who do not need and do not receive accommodations.

Critical Element 5.4 – Monitoring Test Administration for Special Populations

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State monitors test administration in its districts and schools to ensure that appropriate assessments, with or without accommodations, are selected for all students with disabilities and ELs so that they are appropriately included in assessments and receive accommodations that are: • 5.4.1: Consistent with the State's policies for accommodations; • 5.4.2: Appropriate for addressing a student's disability or language needs for each assessment administered; • 5.4.3: Consistent with accommodations provided to the students during instruction and/or practice; • 5.4.4: Consistent with the assessment accommodations identified by a student's IEP Team under IDEA, placement team convened under Section 504; or for students covered by Title II of the ADA, the individual or team designated by a district to make these decisions; or another process for an EL; • 5.4.5: Administered with fidelity to test administration procedures; • 5.4.6: Monitored for administrations of all required academic content assessments and AA-AAAS.	**CE 5.4** •*Evidence#OR1.3.1_2018_19_TAM* •*Evidence#OR1.4.2_2018_19_OAM_Final* •*Evidence#OR1.3.2_OAR_581_022_2100* •*Evidence#OR5.4.1_Division22ComplianceForm* •*Evidence#OR5.4.2_ElemSecondaryComplianceProcess* •*Evidence#OR5.4.3_OAR_581_022_2305* •*Evidence#OR5.4.4_SPR_I_FocusedMonitoring* •*Evidence#OR5.4.5_SPR_I_StateMap* •*Evidence#OR5.4.6_SPR_I_Protocol* •*Evidence#OR5.4.7_SPR_I_and ELaccomm* •*Evidence#OR5.4.8_ELMonitoringTool* •*Evidence#OR5.4.9_ELMonitorTimeline* *Evidence#OR3.1.1.2_Oregon Supplemental Science Peer Review Evidence*, pdf pg. 43 (added by peers)	**CE 5.4:** *Peers found evidence that the State monitors test administration in its districts and schools to ensure that appropriate assessments, with or without accommodations, are selected for all students with disabilities and ELs so that they are appropriately included in assessments and receive accommodations that are consistent with the State's policies for accommodations.* Peers did not find sufficient evidence that the state is monitoring what is actually occurring in schools and districts, but rather is relying on self-reports from districts. While a Focused Monitoring table was provided, there was no description about how a district or school was identified for focused monitoring, or how that process was implemented. The state has plans to implement a system provided by CAI that will provide accommodation usage data on embedded designated supports and accommodations used by students during the assessment. The state will be able to link this usage information to determine the degree of match with IEP, 504, and EL data for students in the system. (*Evidence#OR3.1.1.2_Oregon Supplemental Science Peer Review Evidence*, pdf pg. 43)

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
		It should be noted that the above process will need to be supplemented with continued monitoring of non-embedded supports and accommodations.
Section 5.4 Summary Statement		
___ No additional evidence is required or _X__ The following additional evidence is needed/provide brief rationale: Evidence of the process used by the State to identify sites for focused monitoring.		

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

SECTION 6: ACADEMIC ACHIEVEMENT STANDARDS AND REPORTING

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

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
For academic content standards: The State formally adopted **challenging academic achievement standards** in reading/language arts, mathematics, and science for all students, specifically: • 6.1.1: The State formally adopted academic achievement standards in the required tested grades and, at its option, alternate academic achievement standards for students th the most significant cognitive disabilities; • 6.1.2: The State applies its 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; • 6.1.3: The State's academic achievement standards and, as applicable, alternate academic achievement standards, include: (1) at least three levels of achievement, with two for high achievement and a third for lower achievement; (2) descriptions of the competencies associated with each achievement level; and (3) achievement scores that differentiate among the achievement levels.	**CE 6.1** • *_Evidence#OR1.1.1_OregonStateBoardAuthority* • *_Evidence#OR1.1.2_March6-2014-board-agenda* • *_Evidence#OR1.1.3_March-6-7-2014-draft-minutes* • *_Evidence#OR1.1.4_March6-2014-board-actions* • *_Evidence#OR1.5.1.3_ SB meeting12_12_2019_Science Assessment* • *_Evidence#OR1.5.1.5_12-12-19 State Board Minutes* • *_Evidence#OR1.5.1.4_SB meeting 02_20_2020_Science Approval* • *_Evidence#OR1.5.1.6_2-20-2020 SB Minutes* • *_Evidence#OR1.5.3.2_OSAS Science Assessment_SB Presentation 12_12_2019* • *_Evidence#OR1.5.4.2_V3_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR1.5.2.4_Sci Advisory Aug 7 2018 Agenda* • *_Evidence#OR6.1.1_OSAS Science ALDs* • *_Evidence#OR6.1.2_asmtachstdsummary*	**CE 6.1.1**: *Peers found evidence that the* State *formally adopted academic achievement standards in the required tested grades for science.* **CE 6.1.2:** *Peers found evidence that the* State *applies its 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.* **CE 6.1.3:** *Peers found evidence that the State's academic achievement standards include: (1) at least three levels of achievement, with two for high achievement and a third for lower achievement; (2) descriptions of the competencies associated with each achievement level; and (3) achievement scores that differentiate among the achievement levels.* The state adopted four levels of academic achievement standards – two for high achievement and two for lower achievement.

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
Section 6.1 Summary Statement		
_X__ No additional evidence is required or ___ The following additional evidence is needed/provide brief rationale: • [list additional evidence needed w/brief rationale]		

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 6.2 – Achievement Standards Setting

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State used a technically sound method and process that involved panelists with appropriate experience and expertise for setting: • **Academic achievement standards and, as applicable, alternate academic achievement standards**.	**CE 6.2** •*Please see Common Peer Review Submission for Science Supplemental Evidence and Notes* *Evidence#OR1.5.4.2_V3_OSAS ScienceAssessment Tech Report 2018-2019* •*Evidence#OR1.5.2.4_Sci Advisory Aug 7 2018Agenda* •*Evidence#OR1.5.2.3_2016-2019 Sci AdvisoryMembership Summary*	**CE 6.2:** *Peers found evidence that the State used a technically sound method and process that involved panelists with appropriate experience and expertise for setting academic achievement standards for science.* While the results of the evaluation of the standard-setting workshop generally validated the effectiveness of the workshop, one area of potential concern lies in the lower rankings of agreement to the statement "The Achievement-Level Descriptors (ALDs) (descriptions of what students within each achievement level are expected to know and be able to do) provided a clear picture of expectations for student achievement at each level." (Evidence#OR1.5.4.2_V3_OSAS ScienceAssessment Tech Report 2018-2019, pdf pg 42). Peers recommend that the state consider a follow-up study with teachers from across the state to determine the level of teachers' understanding of the ALDs. For future standard settings, peers recommend increasing the size of the panels to a minimum of 4 to 5 panelists per panel (rather than the 3 panelists from the existing standard setting) and urge the State to continue to seek diverse panelists. Peers suggest that it might be helpful in the future to actually assess teacher panelists on their level of understanding of the NGSS, since frequently teachers report teaching with the NGSS, but lack deep understanding of the standards.

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
Section 6.2 Summary Statement		
_X__ No additional evidence is required or ___ The following additional evidence is needed/provide brief rationale: • [list additional evidence needed w/brief rationale]		

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 6.3 – Challenging and Aligned Academic Achievement Standards

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
For academic achievement standards: 6.3.1: The State's academic achievement standards are challenging and aligned with the State's academic content standards and with entrance requirements for credit-bearing coursework in the system of public higher education in the State and relevant State career and technical education standards such that a 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 they graduate from high school in order to succeed in college and the workforce. 6.3.2: If the State has adopted alternate academic achievement standards for students with the most significant cognitive disabilities, the alternate academic achievement standards (1) are aligned with the State's challenging academic content standards for the grade in which a student is enrolled; (2) promote access to the general curriculum consistent with the IDEA; (3) reflect professional judgment as to the highest possible standards achievable for such students; (4) are designated in the IEP for each student for whom alternate academic achievement standards apply; and (5) are aligned to ensure that a student who meets the alternate academic achievement standards is on track to pursue	**CE 6.3** •*Evidence#OR1.5.4.2_V3_OSAS Science Assessment Tech Report 2018-2019* •*Evidence#OR1.5.2.4_Sci Advisory Aug 7 2018 Agenda* •*Evidence#OR6.1.1_OSAS Science ALDs* •*Evidence#OR1.5.2.3_2016-2019 Sci Advisory Membership Summary* •*Evidence#OR2.1.2.6_V4_OSAS Science Assessment Tech Report 2018-2019*	**CE 6.3.1**: *Peers did **not** find sufficient evidence that the State's academic achievement standards are challenging and aligned with the State's academic content standards and with entrance requirements for credit-bearing coursework in the system of public higher education in the State and relevant State career and technical education standards such that a 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 they graduate from high school in order to succeed in college and the workforce.* Peers did not find evidence that describes the relationship between student academic achievement results and indicators of college or career success. As noted earlier, there seems to be a misalignment between ALD Level 4 and available items on the assessment for students to be able to demonstrate their proficiency at this level. In addition, this observation was noted by some state standard-setting panelists as part of their evaluation comments of the workshop, as documented on pdf pg. 44 of *Evidence#OR1.5.4.2_V3_OSAS ScienceAssessment Tech Report 2018-2019: "*Participants provided suggestions, such as … better aligning the ALDs to the NGSS test items." In other words, the ALDs are designed to be challenging and to be aligned with the NGSS, but the assessment itself appears to lack the highest level of challenge available in the item bank. This may at least partially explain the extremely low percentages of students meeting standard on the assessment at a Level 4. (*Evidence#OR1.5.4.2_V3_OSAS ScienceAssessment Tech Report 2018-2019, pdf pg. 9)* ***CE 6.3.2:** Does not apply.*

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
postsecondary education or competitive integrated employment.		

Section 6.3 Summary Statement

___ No additional evidence is required or

_X__ The following additional evidence is needed/provide brief rationale:

- Evidence that the State's academic achievement standards are challenging and aligned with the State's academic content standards and with entrance requirements for credit-bearing coursework in the system of public higher education in the State and relevant State career and technical education standards.

Critical Element 6.4 – Reporting

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State reports its assessment results for all students assessed, and the reporting facilitates timely, appropriate, credible, and defensible interpretations and uses of those results by parents, educators, State officials, policymakers and other stakeholders, and the public.	**CE 6.4**	
6.4.1: The State reports to the public its assessment results on **student academic achievement for all students and each student group at each achievement level**[3]	*6.4.1* *•Evidence#OR6.4.1.1_State Science Achievement (2018-19)* *•Evidence#OR6.4.1.2_District Science Achievement (2018-19) - Beaverton Example* *•Evidence#OR6.4.1.3_School Science Achievement (2018-19) - Beaverton School Example* *•Evidence#OR6.4.1.4_Assessment Group Report Definitions* *•Evidence#OR6.4.1.5 Public Assessment Group Report (AGR)Webpage Screenshot* *•Evidence#OR6.4.1.6_SchoolDistrictProfiles_Landing* *•Evidence#OR6.4.1.7_Statewide-Annual-Report-Card*	**CE 6.4.1** *Peers found sufficient evidence that the state reports to the public its assessment results on student academic achievement for all students and each student group at each achievement level.*

[3] Although all students with disabilities must be included in a State's assessment system, requirements for public reporting in ESEA section 1111(h)(1)(C)(ii) apply only to children with disabilities as defined in section 602(3) of the IDEA.

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
6.4.2 For **academic content assessments**, 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 academic content assessments that: o Provide valid and reliable information regarding a **student's**	***6.4.2*** • *_Evidence#OR4.7.1_V6_OSAS Science Assessment Tech Report 2018-2019* • *_Evidence#OR6.4.2.1_osas_reports_userguide* • *_Evidence#OR6.4.2.2_19_20_OSASToolkit_Sci_Dist_Info* • *_Evidence#OR6.4.2.3_19_20_OSAS_Sci_Toolkit_Parent* • *_Evidence#OR6.4.2.4_9_20_OSAS_Sci_Toolkit_Parent_SPANISH* • *_Evidence#OR6.4.2.5_19_20_OSASToolkit_Sci_Teacher* • *_Evidence#OR6.4.2.6_Mockup_ISR_NewScience* • *_Evidence#OR6.4.2.7_Mockup_CombinedISR_NewScience* • *_Evidence#OR6.4.2.8_ParentISROtherEngSpan*	**CE 6.4.2** *Peers found **partial** evidence to support this CE. Specific comments regarding the bulleted items under 6.4.2 are noted below:* o With regard to providing valid and reliable information regarding a student's academic achievement: Because the results are not reported in terms of all three NGSS dimensions (SEPs, CCCs, and DCIs), the interpretation of results in light of instruction and policy-making will be limited. However, the document "Understanding and Preparing for the New More Rigorous OSAS Science Assessment" uses the following wording: that DCIs, SEPs, and CCCs are woven together in the items within the assessment and "are not reported separately at this time." Hopefully the state does plan to make adjustments to the reporting to address this important issue regarding optimal use of the results (*Evidence#OR6.4.2.2_19_20_OSASToolkit_Sci_Dist_Info)* o *Peers found evidence that the state reports the student's academic achievement in terms of the State's grade-level academic achievement standards.* o With regard to the state providing information to help parents, teachers, and principals interpret the test results and address the specific academic needs of students: Peers would like to see standard error (e.g., a confidence band) included in student reports, to indicate to parents and other stakeholders that there is a degree of measurement error inherent in score reporting. o *Peers found evidence that reports are provided in an understandable and uniform format.* o *Peers did **not** find evidence that reports a*re, *to the extent practicable, written in a language that parents and*

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
academic achievement; o Report the **student's academic achievement** in terms of the State's grade-level academic achievement standards; o Provide information to help parents, teachers, and principals interpret the test results and address the specific **academic needs of students**; o Are provided in an understandable and uniform format; o Are, to the extent practicable, written in a language that parents and guardians can understand or, if it is not		*guardians can understand or, if it is not practicable to provide written translations to a parent or guardian with limited English proficiency, are orally translated for such parent or guardian.* Specifically, peers did not find evidence that reports are written in a language other than English or Spanish. The State's submission indicates that districts are responsible for this, so peers request evidence of supports being provided to districts and monitoring to ensure that districts are meeting this requirement appropriately. o With regard to providing reports in an alternative format accessible to parents with a disability as defined by ADA, the State's submission indicates that districts are responsible for providing such reports. Peers request evidence of supports being provided by the state to districts for meeting this responsibility and monitoring to ensure that districts are meeting this requirement appropriately.

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
practicable to provide written translations to a parent or guardian with limited English proficiency, are orally translated for such parent or guardian; o Upon request by a parent who is an individual with a disability as defined by the ADA, as amended, are provided in an alternative format accessible to that parent.		
6.4.3 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.	***6.4.3*** • *_Evidence#OR2.1.1.2_OAR581_022_2270* • *_Evidence#OR6.4.3.1_AsmtAcctbltyChecklist1920*	**CE 6.4.3**: *Peers **did not** find sufficient evidence that 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.* Though this indicator of timeliness was disrupted by COVID-19, panelists would like to see evidence/proof points of the timeliness of report delivery in normal (non-pandemic) times.

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence

Section 6.4 Summary Statement

___ No additional evidence is required or

_X__ The following additional evidence is needed/provide brief rationale:

- Evidence that, once the assessment is operational on a normal schedule (no COVID-19) stakeholders receive reports in a timely manner.
- Evidence that reports are, upon request by a parent who is an individual with a disability as defined by the ADA, as amended, are provided in an alternative format accessible to that parent.

SECTION 7: LOCALLY SELECTED NATIONALLY RECOGNIZED HIGH SCHOOL ACADEMIC ASSESSMENTS

(**if applicable**; evidence for this section would be submitted in ADDITION to evidence for sections 1 through 6)

Critical Element 7.1 – State Procedures for the Use of Locally Selected, Nationally Recognized High School Academic Assessments

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State has established technical criteria to use in its review of any submission of a locally selected, nationally recognized high school academic assessment. The State has completed this review using its established technical criteria and has found the assessment meets its criteria **prior to** submitting for the Department's assessment peer review. The State's technical criteria include a determination that the assessment: • Is aligned with the challenging State academic standards; and • Addresses the depth and breadth of those standards. AND The State has procedures in place to ensure that a district that chooses to use a nationally recognized high school academic assessment administers the same assessment to all high school students in the district except for students with the most significant		

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
cognitive disabilities who may be assessed with an AA-AAAS. AND The technical criteria established by the State in reviewing a locally selected, nationally recognized high school academic assessment must ensure that the use of appropriate accommodations does not deny a student with a disability or an EL— • The opportunity to participate in the assessment; and • Any of the benefits from participation in the assessment that are afforded to students without disabilities or students who are not ELs.		

Section 7.1 Summary Statement

___ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:
- [list additional evidence needed w/brief rationale]

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Element 7.2 –State Monitoring of Districts Regarding the Use of Locally Selected, Nationally Recognized High School Academic Assessments

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State must have procedures in place to ensure that: **Before** a district requests approval from the State to use a nationally recognized high school academic assessment, the district notifies all parents of high school students it serves— • That the district intends to request approval from the State to use a nationally recognized high school academic assessment in place of the statewide academic assessment; • Of how parents and, as appropriate, students may provide meaningful input regarding the district's request (includes students in public charter schools who would be included in such assessments); and • Of any effect of such request on the instructional program in the district.		

Section 7.2 Summary Statement

___ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:
- [list additional evidence needed w/brief rationale]

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

STATE ASSESSMENT PEER REVIEW NOTES FOR Oregon - Science

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Element 7.3 –Comparability of the Locally Selected Nationally Recognized High School Academic Assessments with the State Assessments

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The locally selected, nationally recognized high school academic assessment: • Is equivalent to or more rigorous than the statewide assessment, with respect to— o The coverage of academic content; o The difficulty of the assessment; o The overall quality of the assessment; and o Any other aspects of the assessment that the State may establish in its technical criteria; • Produces valid and reliable data on student academic achievement with respect to all high school students and each subgroup of high school students in the district that— o Are comparable to student academic achievement data for all high school students and each subgroup of high school students produced by the statewide assessment at each academic achievement level; o Are expressed in terms consistent with the State's academic achievement standards; and o Provide unbiased, rational, and consistent differentiation among schools within the State for the purpose of the State determined accountability system including calculating the Academic Achievement indicator and annually meaningfully differentiating between schools.		

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
Section 7.3 Summary Statement		
___ No additional evidence is required or ___ The following additional evidence is needed/provide brief rationale: • [list additional evidence needed w/brief rationale]		

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

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

July-August 2020 State Assessment Peer Review Notes

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

Note: Peer review notes provide the combined recommendations of the individual peers to the U.S. Department of Education (Department), based on the statute and regulations, the Department's peer review guidance, and the peers' professional judgement of the evidence submitted by the State. These assessment peer review notes, however, do not necessarily reflect 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 for assessment peer review. Although the peer notes inform the Secretary's consideration of each State's assessment system, the Department makes the final decision regarding whether the assessment system meets the requirements in the statute and regulations. As a result, these peer notes may not completely align with the final determination made by the Department.

Contents

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

SECTION 1: STATEWIDE SYSTEM OF STANDARDS AND ASSESSMENTS

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

Critical Element	**Evidence (Record document and page # for future reference)**	**Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence**
For academic content standards: 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 schools and public school students in the State.	See State-Specific Submission	This critical element was met in a 2019 peer review for reading/language arts (R/LA), mathematics, and science.

Section 1.1 Summary Statement

_X__ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:

- [list additional evidence needed w/brief rationale]

Critical Element 1.2 – Challenging Academic Content Standards

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
For academic content standards: The State's challenging academic content standards in reading/language arts, mathematics, and science are aligned with entrance requirements for credit-bearing coursework in the system of public higher education in the State and relevant State career and technical education standards.	See State-Specific Submission	This critical element was met in a 2019 peer review for R/LA, mathematics, and science.

Section 1.2 Summary Statement

_X__ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:

- [list additional evidence needed w/brief rationale]

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 1.3 – Required Assessments

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State's assessment system includes annual general and alternate assessments aligned with **grade-level academic achievement standards** or alternate academic achievement standards in: • Reading/language arts (R/LA) and mathematics in each of grades 3-8 and at least once in high school (grades 9-12); • Science at least once in each of three grade spans (3-5, 6-9 and 10-12). AND The State's **academic content assessments** must be the same assessments administered to all students in the tested grades, with the following exceptions: • Students with the most significant cognitive disabilities may take an alternate assessment aligned with alternate academic achievement standards. • A State may permit an LEA to administer a nationally recognized high school academic assessment in lieu of the State high school assessment if certain conditions are met. • A State that administers an end-of-course high school mathematics assessment may exempt an 8th grade student from the mathematics assessment typically administered in	Reviewed by Department Staff Only	This critical element was met in a 2019 peer review for R/LA, mathematics, and science.

eighth grade and allow the student to take the State end-of-course mathematics test instead. • The Department may have approved the State, under the Innovative Assessment Demonstration Authority, to permit students in some LEAs to participate in a demonstration assessment system in lieu of participating in the State assessment.		

Section 1.3 Summary Statement

_X__ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:

- [list additional evidence needed w/brief rationale]

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 1.4 – Policies for Including All Students in Assessments

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or 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 those children 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 ELs: o Policies state that all ELs must be included in all aspects of the content assessment system, unless the State has chosen the statutory option for recently arrived ELs under which such ELs are exempt from one administration of its reading/ language arts assessment. o If a State has developed native language assessments for ELs in R/LA, ELs must be assessed in R/LA 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	Reviewed by Department Staff Only	This critical element was met in a 2019 peer review for R/LA, mathematics, and science.

language assessments for a period not to exceed two additional consecutive years. ○ If the State uses the flexibility for Native American language schools and programs: (1) the State provides the content assessment in the Native American language to all students in the school or program; (2) the State submits such content assessment for peer review as part of its State assessment system; and (3) the State continues to provide ELP assessments and services for ELs as required by law. The State must assess in English the students' achievement in R/LA in high school.		

Section 1.4 Summary Statement

_X__ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:

- [list additional evidence needed w/brief rationale]

Critical Element 1.5 – Meaningful Consultation in the Development of Challenging State Standards and Assessments

(Note: this is a new requirement under ESSA, so it does not apply to standards and assessments adopted prior to the passage of ESSA (December 2015)).

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
If the State has developed or amended challenging **academic** standards and assessments, the State has conducted meaningful and timely consultation with: • State leaders, including the Governor, members of the State legislature and State board of education (if the State has a State board of education). • Local educational agencies (including those located in rural areas). • Representatives of Indian tribes located in the State. • Teachers, principals, other school leaders, charter school leaders (if the State has charter schools), specialized instructional support personnel, paraprofessionals, administrators, other staff, and parents.	Reviewed by Department Staff Only	This critical element was met in a 2019 peer review for R/LA and mathematics. The State adopted science standards in April 2015, prior to the passage of ESSA (December 2015).

Section 1.5 Summary Statement

_X__ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:

- [list additional evidence needed w/brief rationale]

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

SECTION 2: ASSESSMENT SYSTEM OPERATIONS
Critical Element 2.1 – Test Design and Development

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State's test design and test development process is well-suited for the content, is technically sound, aligns the assessments to **the depth and breadth of the State's academic content standards** for the grade that is being assessed and includes:	*CS 001 – V1_Shared Technical Report.pdf*	While the Cambium Assessment, Inc. (CAI) consortium has developed an item bank to support each participating state's development of a science assessment and has provided guidance to states on the development of an assessment that will measure the depth and breadth of each state's standards, it is the **role of each participating state** to: • **Adopt their version of science standards, which has either been the Next Generation Science Standards (NGSS) or a related set of standards based on the same 3-dimensional framework.** • **Develop test blueprints that align to their state standards.** • **Implement a test design to support their intended purposes.** • **Submit evidence demonstrating how the guidance from Cambium has been operationalized in each state's customized assessment**.
• (CE 2.1.1) Statement(s) of the purposes of the assessments and the intended interpretations and uses of results;	CE 2.1.1: ● *CS 001 – V1_Shared Technical Report.pdf* o Section 1.2 - outlines the typical purposes of statewide assessments and the types of assessments that the Shared Science Assessment Item Bank is designed to support (pp. 2-3; pdf pp. 7-8) o Section 1.2 - Describes the unique features of three-dimensional science standards and the types of interpretations about student performance that appropriately aligned assessments can support (p. 2; pdf p. 7)	● CE 2.1.1: CAI has described one purpose of the assessment – to measure each student's performances against a multi-dimensional science standards framework. CS 002, p.2 states "The goals, uses, and claims the Shared Science Assessment Item Bank and subsequent tests would be designed to support were identified in a series of collaborative meetings held August 22–23, 2016." However a full statement of these goals, uses, and claims is not found in the evidence. Statement of the intended interpretations of results is implicit rather than explicit. Evidence from the

	• *CS 002 – V2_Shared Technical Report, p. 2* • *CS 004.C – Alignment Report*, p.54.	independent alignment study (*CS 004.C – Alignment Report*, p.54) summarizes results of analyses showing support for "reasonable inferences" about a student's proficiency based on the three-dimensional expectations within the standards. However, given that the results are not reported in terms of the three dimensions, it is dubious if users will be able to make interpretations of the results in terms of the three dimensions of the standards. CAI has provided a basic list of intended uses of the results (page 3 of *CS 001*), such as an indicator of academic achievement and progress aligned to academic content standards, but the list includes bulleted elements that are simply actions related to test administration, scoring, and reporting, without regard to actual uses of the results. Also it must be noted that no evidence is found to support the assessment being an indicator of student progress or growth over time. In addition, there is no mention of uses related to instructional purposes by teachers or self-assessment by students. **States will need to indicate if they intend to meet instructionally-related uses.**
(CE 2.1.2) Test blueprints that describe the structure of each assessment in sufficient detail to support the development of assessments that are technically sound, measure the depth and breadth of **the State's grade-level academic content standards** and support the intended interpretations and uses of the results.	CE 2.1.2: ● *CS 002 – V2_Shared Technical Report.pdf.pdf* ○ Section 1.1 – Identifies the structure of three-dimensional science standards based on NRC's framework (e.g., NGSS) as the claim structure the Shared Science Assessment Item Bank is designed to support (p. 2; pdf p. 5) ○ Section 3.1 – Describes the current composition of the Shared Science Assessment Item Bank (pp. 19-24; pdf pp. 22-27) ○ Section 4.1 – Provides a description of test design and blueprint development processes (pp. 25-26; pdf p. 28-29)	CE 2.1.2: CAI describes guidance they provided to states in the development of test blueprints, including training on the multi-dimensional nature of the standards, how items were constructed to align with the standards, steps to follow, and a detailed listing of the components of each blueprint (*CS 002 J,* Section 1.1) as well as general guidelines to be followed by the states (*CS 002-V2*, Section 4.2). However, while general guidelines and components for the blueprints are provided, it would be helpful for CAI to provide the specific rules, constraints, and parameters for building the blueprints to fully document the technical soundness of the process. Similarly it would be helpful for

	○ Section 4.2 – Describes the expected characteristics of test blueprints (p. 26; pdf p. 29) ● *CS 002.J – Appendix J. Adaptive Algorithm Design.pdf* ○ Section 1.1 – Details expected blueprint components and constraints (pp. 5-6)	CAI to provide a detailed explanation of their procedure for reviewing each state's blueprints. *CS 002.J – Appendix J,* p. J-4, mentions that an example of a blueprint specification is included in Appendix J-1, but it is not found. The inclusion of a sample blueprint would be very helpful to further document the technical soundness of the process. *CS 002 – V2*, p. 26 states that "Blueprints (i.e., test specifications) report the expected numbers of items by category (e.g., stand-alone, item cluster) for each science discipline or Disciplinary Core Idea (DCI)." This and additional evidence indicate that participating states can construct their test blueprints based on DCI coverage alone, without regard to the other two dimensions – Science and Engineering Practices (SEPs) and Cross-Cutting Concepts (CCCs). No evidence is found to document how the blueprints will be constructed to reflect an appropriate balance of all three dimensions of the standards. To properly meet the breadth and depth of the standards, developers need to provide ways to ensure that participating states have blueprints that allow users to evaluate students' use of the eight SEPs as well as which clusters/items include student use of the CCCs to demonstrate sense-making. **States will need to document how they followed the guidelines provided by CAI and include evidence of their test blueprints meeting the required components at an acceptable level.**
• (CE 2.1.3) Processes to ensure that each academic 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	CE 2.1.3: ● *CS 002 – V2_Shared Technical Report.pdf* ○ Section 1.1 - Identifies the structure of three-dimensional science standards based on NRC's framework (e.g., NGSS) as the claim structure	CE 2.1.3: Because CAI developed the item bank using item clusters that encompass all three dimensions represented in the standards (in the item specifications), and because the item types vary in terms of how students demonstrate their performances (*CS 002, V2*, Section 3.1), thus providing

applications of knowledge and skills (i.e., higher-order thinking skills).	the Shared Science Assessment Item Bank is designed to support (p. 2; pdf p. 5) ○ Section 3 - Summarizes the item bank (pp. 18-25; pdf pp. 21-28) ○ Section 4.1 – Provides a high-level discussion of how CAI works with participating states to develop assessments that meet the states' blueprint requirements (p. 25-26; pdf pp. 29-30) • *CS 004.A – Appendix A. Science Clusters Cognitive Lab Report.pdf* ○ Section 3.1.1 – Identifies improvements for specific item types (pp. 6-7) ○ Section 3.5.2 – Describes students' use of similar online tests and tools (p. 102) • *CS 004.B – Braille Cognitive Lab Report.pdf* ○ Section 4 – Describes specific recommendations emerging from the braille cognitive lab (pp. 8-9)	flexibility in student responses, the test development process does allow for appropriate inclusion of challenging content. In addition, because the items within each cluster build a cognitive sequence, the assessment then requires complex demonstrations of applications of knowledge and skills in order to complete the items within each cluster. The Cognitive Lab study that was conducted provides student-level evidence of the complexity in thinking required in the assessment. The fact that the students included were "high achievers" in science but found the assessment to be of "medium" difficulty, and the students' descriptions of their thinking strategies, supports that the assessment requires higher-level thinking. While the Cognitive Lab study conducted with visually impaired students surfaced a number of technological and formatting issues that need to be addressed, this issue relates more to the validity of the assessment results for this sub-population than to the appropriateness and challenging nature of the content itself. Based on the visually-impaired students' description of the strategies and thinking processes they used, the content per se appeared to be challenging. In the documentation provided by CAI, it would be helpful when using the term "skills" to differentiate between skills used for accessing the technology and skills used in demonstrating proficiency in the Science and Engineering Practices. Although the item specifications include all three dimensions of the standards, as mentioned above evidence indicates the likelihood of states developing blueprints based on content (DCI) alone. Evidence is not found that would offer guidance to states regarding how to represent the other two dimensions of the standards in their blueprints. The three-dimensionality of the NGSS framework is designed to ensure that science achievement

• (CE 2.1.4) If the State administers computer-adaptive assessments, the item pool and item selection procedures adequately support the test design and intended uses and interpretations of results.	CE 2.1.4: ● *CS 002 – V2_Shared Technical Report.pdf* o Section 5 – Describes evaluation of item exposure using simulation summaries (p. 27-28; pdf pp. 31-32) o Section 3.2 – Discusses future item development plans to refresh item pool as item exposure increases (p. 25; pdf p. 29) ● *CS 002.J − Appendix J. Adaptive Algorithm Design.pdf* o Provides an overview of the item selection algorithm supported by CAI's assessment delivery platform	reflects the ability to make sense of challenging content and to perform complex demonstrations and applications of knowledge and skills. If the emphasis in this assessment is on content (what they know) alone, there is no assurance that students' higher-order thinking skills are being elicited and that they are able to demonstrate what they are able to do with that content. It is important for CAI to include guidance to states on how to build blueprints to include SEPs and CCCs. CE 2.1.4: CAI has constructed this assessment so that it can be given in fixed or adaptive form, and in segmented (all items within the same discipline) or unsegmented format. The documents include information on the selection algorithm and the relationship between the item pool and the constraints to be met. All the participating states in 2019 chose to use a LOFT (linear-on-the-fly) design, which allows for students to see different sets of items, although the selection of items is not based on the performance of the student. Rather the LOFT design varies the items that a student will see based on parameters designed to ensure that the test blueprint is maintained. It appears that the LOFT design has the advantage of students not all seeing the same items on a test, but it lacks the advantage of an adaptive test that would optimize the number of items a student sees and therefore the amount of time needed to zero in on a student's functioning level. CAI plans to refresh and expand the item pool to allow computer adaptive testing for all participating states in the future. There is evidence that there are procedures in place for replenishing the item pool, for evaluating overused items and tracking content coverage (i.e., DCI only).

• (CE 2.1.5) If the State administers a computer-adaptive assessment, it makes proficiency determinations with respect to the grade in which the student is enrolled and uses that determination for all reporting. • (CE 2.1.6) If the State administers a content assessment that includes portfolios, such assessment may be partially administered through a portfolio but may not be *entirely* administered through a portfolio.	CE 2.1.5: ● *CS 002.J – Appendix J. Adaptive Algorithm Design.pdf* o Describes the process by which the item selection algorithm selects items within the appropriate grade band for each test event ● *CS 006 − V6_Shared Technical Report.pdf* o Section 1 - Describes the score reports available through CAI's score reporting system (pp. 1-2; pdf pp. 3-4) o Section 2 - Overviews the interpretation of typical reported scores and shows that all scores are reported at grade-level (p. 2; pdf p. 4) CE 2.1.6: Not applicable	The processes used in the development and implementation of the LOFT approach appear to adequately support the test design in an overall sense. **The states will need to document that the use of the LOFT design with their items did in fact support their test design, and their intended uses and interpretations of results. The state documentation should include item exposure rates for the spring 2019 operational administration** (*CS 002 – v2*, pdf pg. 27). CE 2.1.5: The computer adaptive algorithm used in this assessment selects only on-grade level items that contribute to a student's overall score. Participating states may choose to customize their score reports for their state programs, but all states report scores at grade-level as determined by items within the applicable grade band. Appendix J has a strong explanation of the algorithm which seems appropriate, with a focus on proficiency determination and sub-score reporting. "Cambium Assessment, Inc. (CAI) envisions … very high or very low achieving students to continue on to a segment that contains items from adjacent grades but barring other students from those segments" (p. J-16). This seems to be in line with the requirement to first test for proficiency and then find the location. CE 2.1.6: Not applicable
Section 2.1 Summary Statement		
___ No additional evidence is required or _X_ The following additional evidence is needed/provide brief rationale: • CE 2.1.1: Provide an explicit explanation of the purposes and interpretations of the use of the results.		

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

- CE 2.1.2: Provide the specific rules, constraints, and parameters for building the blueprints as well as a detailed explanation of the procedure for reviewing each state's blueprints.
- CE 2.1.2: Provide evidence that the test blueprints that will be developed by states will support the development of assessments that measure the depth and breadth of the State's grade-level academic content standards, and support the intended interpretations and uses of the results, in light of the three dimensions of the academic content standards adopted by each state and the absence of two of these dimensions in the proposed blueprint designs.
- CE 2.1.3: Provide evidence that the processes used in test design and development include guidance for including SEPs and CCCs throughout the process (e.g., blueprints, scoring, results) to ensure that each academic assessment is tailored to the knowledge and skills included in the three-dimensional academic content standards.

Critical Element 2.2 – Item Development

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or 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.	● *CS 002 − V2_Shared Technical Report.pdf* o Section 1.1 − Identifies the structure of three-dimensional science standards (p. 2; pdf p. 5) o Section 1.2 – Describes the process for ensuring proper item review steps, and the underlying principles guiding development (pp. 2-3; pdf pp. 5-6) ▪ Evidence-centered design ▪ Universal design ▪ Accessibility (WCAG) 2.0 o Section 2.1 (Exhibit A) − Overviews the item development process to support the validity of claims and each state's involvement (p. 4; pdf p. 7) o Section 2.2 – Presents the item specifications and sample item clusters that guide item-writing and review work (pp. 4-7; pdf pp. 7-10) o Section 2.3 – Describes qualifications and training of item writers (p. 7; pdf p. 10) o Section 2.4 – Details internal review processes (pp. 7-10; pdf pp. 10-13) o Section 2.5 – Details participating state review processes (pp. 10-13; pdf pp. 13-17) o Section 2.6-2.7 – Describe field testing and data review procedures (pp. 14-18; pdf pp. 17-21) o Section 2.7.1 – Provides the rubric validation process (pp. 14-15; pdf pp. 17-18) o Section 3 – Describes the item types within the item bank (pp. 18-15; pdf pp. 21-28) ● *CS 002.A – Appendix A. Item Writer Training Materials.pdf* o Provides the power point outlining key item writing materials, including Language	The process used by CAI in item development appears to be comprehensive and robust, including training for item writers and steps for item review, and a comprehensive field-test analysis of items, including Differential Item Functioning (DIF) and Item Response Theory (IRT). All items were reviewed for potential bias, sensitivity, and accessibility issues. The item specifications were quite detailed, supporting the item development process in validly measuring the three-dimensional science standards in both content and in the continuum of cognition needed by students to perform successfully (including SEPs and CCCs as well as DCIs). The item review process was conducted on all the items that became part of the shared item bank, including both the CAI shared (common) items and the state-specific items that were included in the item bank for use by other participating states. Despite the fact that the item specifications included all three dimensions, the focus was clearly on the content, with every item required to have a DCI component but only one of the other two dimensions. While this approach might seem reasonable in a traditional paradigm, it appears to have limited the test developers in their ability to embed all three dimensions throughout the process following item development. In other words, it appears that a traditional one-dimensional approach was used to assess an innovative three-dimensional model. An assessment built on a three-dimensional framework such as NGSS requires additional considerations to fully capture the nature of the standards being assessed. Evidence was not found showing how the traditional approach has been modified/customized or will

	Accessibility, Bias, and Sensitivity Guidelines and item-writer training materials ● *CS 002.B – Appendix B. Item Review Checklist.pdf* o Provides the item checklist ● *CS 002.C – Appendix C. Content Advisory Committee Participant Details* ● *CS 002.D – Appendix D. Fairness Committee Participant Details* ● *CS 002.H – Appendix H. Science Item Bank.pdf* o Shows item bank details by grade band, performance expectation, and origin ● *CS 002.K – Appendix K. Item Specifications – Grade 3-High School.pdf* o Describes item specifications by grade level for each Performance Expectations. o Identifies content and content limits and task demands to provide specific guidance to item writers ● *CS 002.L – Appendix L. Style Guide for Science Items.pdf* o Provides guidance for item writers to ensure item presentation is consistent, accessible, and appropriate for the mode of test administration ● *CS 001 – V1_Shared Technical Report.pdf* o Section 3.1.1 – Describes the 2018 field test procedures (initial field testing for the Shared Science Assessment Item Bank) (pp. 8-13; pdf pp. 13-18) o Section 3.1.2 – Describes the 2019 field testing procedures (pp. 14-18; pdf pp. 19-23)	be modified/customized in the future so that it can truly assess a three-dimensional framework. In addition to the items developed for the Shared Science Assessment Item Bank, **participating states may require custom-developed items to sufficiently address unique blueprint requirements or address state-specific content standards. In such cases, additional item development and selection processes will need to be documented in the individual state's Peer Review submission.** Within document *CS 002.K – Appendix K. Item Specifications – Grade 3-High School,* there seems to be a discrepancy between expected vocabulary listed and the information provided within the Framework and the NGSS. For example 3-PS2-2 (page K-7) indicates students should know vocabulary including mass, electrical field, charged particles. According to the DCI progressions: • The concept 'fields' is not introduced until the middle school grade band. • Until middle school, the term 'weight' is used as the 5th grade DCI indicates: '…does not distinguish between mass and weight.' • The concept 'particles' does not appear until 5th grade. • "Charged particles" is not identified until high school. • Although the word 'gravity' is identified in the assessment boundary, learning around gravity is not specifically identified until 5th grade. MS-LS1-3 is another example of unnecessary expected vocabulary (e.g., aorta, artery, bone marrow, and several organs). Although these words could be used within lessons, neither the PE nor the DCI indicate that a specific body system or systems need to be included in student learning. In addition, identifying specific 'expected to know' vocabulary could be seen as guiding teachers to address a

		specific PE in a certain way or lead them to use certain curricular materials. Identifying assessment-specific vocabulary could also lead to item creation that shows bias to students with special needs and EL students. *CS 002-B –Item Review Checklist,* (p B-1 & B-2) refers to checking that the reading level, language and vocabulary as well as mathematical elements such as graphs and tables are appropriate for the grade level. It is not clear how the item writers were to determine grade-level appropriateness, whether using Common Core Standards or other references. *CS 002 – V2*, Section 2.3, states that item writers have at least a BA degree and might have teaching experience. It is unclear (a) how the item writers were recruited and (b) their level of expertise in terms of science, specifically their experience and understanding of the NGSS. In addition, item writers should represent a spectrum of expertise across the science disciplines. Future training for item writers would be strengthened by including: (a) comprehensive training regarding the meaning and use of the NGSS, (b) examples of quality items for the writers to have models of well-written items, and (c) checkpoint procedures and coaching as needed during the training and the item-writing process to ensure that the item writers understand the requirements and expectations for quality items that reflect three-dimensional science (e.g., having the writers identify flaws in sample items and explain how the items could be improved). It would be helpful to know the percentage of items created by the item writers in the first phase needed to be modified or completely rewritten by subsequent reviewers. If a substantial portion of items need to be rewritten by "senior" staff, then the initial item writing may not truly be providing opportunity for input at the teacher level.

		The lists of teachers on the Content Advisory Committees (*CS 002-C*) and the Fairness Committees (*CS 002-D*) were incomplete in terms of providing participant demographics and did not specify NGSS/science expertise for the participants. In addition, most of the groups were rather small in size when divided by grade band. From "*CS 002.B - Appendix B. Item Review Checklist.pdf*": • Claiming "Science vocabulary should be part of the 'Science Vocabulary Students Are Expected to Know' in the item specifications" is probably appropriate, but implies a potential focus on recall of terms. It would be helpful for CAI to clarify the use of the Vocabulary terms in the item specifications. • Some checklist items (e.g., "Do the interactions avoid redundancy?") seem complicated and/or vague. It would be helpful for CAI to do some type of check of item writers' and reviewers' understanding of the checklist items prior to their using this checklist.

Section 2.2 Summary Statement

___ No additional evidence is required or

X_ The following additional evidence is needed/provide brief rationale:

- Provide evidence that the item development process is appropriate to sufficiently develop quality items that assess a three-dimensional framework of standards, including greater details about:
  - ways that the process can fully reflect all three dimensions
  - characteristics of item writers and other committee review members, especially in terms of science/NGSS expertise
  - discrepancies between NGSS and the item specifications
  - the purpose of vocabulary terms being called out in the item specifications).

Critical Element 2.3 – Test Administration

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State implements policies and procedures for standardized test administration; specifically, the State: • (CE 2.3.1) Has established and communicates to educators clear, thorough and consistent standardized procedures for the administration of its assessments, including administration with accommodations;	CE 2.3.1: ● *CS 005 – V5_Shared Technical Report.pdf* o Section 2 - Provides standardized administration procedures (pp. 2-4; pdf pp. 4-6) o Section 5 - Outlines online features and accommodations including EL accommodations (pp. 7-11; pdf pp. 9-13)	CE 2.3.1: CAI provided to participating states (a) an online interactive Test Administrator (TA) Certification Course designed to help TAs know the steps for administering a test session in the online system, and (b) a practice test for TAs and students, as well as (c) multiple online training opportunities for key staff. (pg. 2-3) CAI also provided manuals and user guides related to the online technical specifications and user guides related to navigating various aspects of the online system. (pg. 3) The range of embedded and non-embedded universal tools, supports, and accommodations built into the system and allowed for designated students are inclusive of features available in the most widely used state assessments. (pg. 7-13). CAI provided manuals that address implementing the Braille version of the test, supports recommended for English Learners (ELs), and how to digitally manage student settings regarding accommodations. **States will need to provide evidence regarding**: • **TA training provided, including the state TA manual and other documentation provided to TAs in order for them to follow the standardized administration procedures.** • **TA completion of the TA Certification Test.** • **Requirements for students to take the practice test. (CS 004-v4, p. 20)** • **A description of available accessibility features, supports, and accommodations used in their state.** • **Identifying digitally-embedded and non-embedded designated supports for students for whom an adult or team has indicated a need for the support. (p. 9)**

• (CE 2.3.2) Has established procedures to ensure that general and special education teachers, paraprofessionals, teachers of ELs, specialized instructional support personnel, and other appropriate staff receive necessary training to administer assessments and know how to administer assessments, including, as necessary, alternate assessments, and know how to make use of appropriate accommodations during assessments for all students with disabilities; • (CE 2.3.3) 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.	CE 2.3.2: See State-Specific Submissions (CE 2.3.3) See State-Specific Submissions	CE 2.3.2: See State-Specific Submissions (CE 2.3.3) Evidence was not found regarding contingency plans for server difficulties that originate at the CAI delivery level. **See State-Specific Submissions for addressing this element at the local level.**

Section 2.3 Summary Statement

___ No additional evidence is required or

_X_ The following additional evidence is needed/provide brief rationale:

- Evidence regarding contingency plans for server difficulties that originate at the CAI delivery level.

Critical Element 2.4 – Monitoring Test Administration

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or 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. Monitoring of test administration should be demonstrated for all assessments in the State system: the general academic assessments and the AA-AAAS.	Reviewed by Department Staff Only	Department staff determined that the State's evidence is sufficient for this critical element.

Section 2.4 Summary Statement

_X__ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:

- [list additional evidence needed w/brief rationale]

Critical Element 2.5 – Test Security

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or 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:		The common submission describes the security procedures included in developing test content, as well as features of the assessment delivery system that ensure test security. **Participating states' submissions will need to include documentation of specific test security procedures (e.g., training for testing personnel, reporting of incidents, consequences) in their individual state submissions.**
• (CE 2.5.1) Prevention of any assessment irregularities, including maintaining the security of test materials (both during test development and at time of test administration), 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;	CE 2.5.1: ● *CS 005 − V5_Shared Technical Report.pdf* o Section 3.2 − Provides information regarding a secure browser and system security (pp. 5-6; pdf pp. 7-8) o Section 3.2 − Discusses prevention of test irregularities via a secure test environment (pp. 5-6; pdf pp. 7-8) o Section 3.4 − Provides preparation guidelines for test irregularities and other security violations (p. 6; pdf p. 8) ● *CS 002.E – Appendix E. Sample Data Review Training Materials.pdf* o Provides an example of non-disclosure requirement (p. 2) ● *CS 003.B – Appendix B. Standard-Setting Training Slides.pdf* o Example of security discussion during standard-setting training (p. 20)	CE 2.5.1: During test development training, CAI required item writers to sign non-disclosure forms. (*CE 002-E*, p. 2) During standard setting, CAI established procedures that protected the security of the items being viewed (*CS 003-B*, p. 20). CAI has put into place several mechanisms to avoid assessment irregularities related to test security during test administration, including setting up the online testing system (TIDE) with staff having specific roles in terms of level of access to secure information, and developing a secure browser that prevents students from accessing other browsers or apps during testing.
• (CE 2.5.2) Detection of test irregularities;	CE 2.5.2: ● *CS 005 – V5_Shared Technical Report.pdf* o Section 3.4 – Test Security Violations (p. 27)	CE 2.5.2: Evidence was not found regarding procedures set up by CAI for the detection of test irregularities in terms of potential loss of security over specific test items, especially for those that might occur during test development or standard-setting that is under the control of CAI. The

		description on Page 1 of *CS 005* says participant states will describe detection of test irregularities, but Cambium ought to be running those checks as well. Because CAI has cross-state data, they are best suited to detect anomalies and also to recognize when any items have been exposed. **States need to provide evidence that they have procedures in place for detecting test irregularities that occur during any item development within the state as well as during test administration that might compromise the security of test materials.**
• (CE 2.5.3) Remediation following any test security incidents involving any of the State's assessments; • (CE 2.5.4) Investigation of alleged or factual test irregularities. • (CE 2.5.5) Application of test security procedures to all assessments in the State system: the general academic assessments and the AA-AAAS.	CE 2.5.3: CE 2.5.4: See State-Specific Submissions. CE 2.5.5: See State-Specific Submissions.	CE 2.5.3: No evidence is found regarding procedures set up by CAI for follow-up or remediation for test irregularities that might occur in terms of loss of security over specific test items during item development, standard-setting or other test development activities that are under the control of CAI. **States need to provide evidence that they have procedures in place for remediating test irregularities that occur during any item development within the state as well as during test administration.**

Section 2.5 Summary Statement

___ No additional evidence is required or

_X_ The following additional evidence is needed/provide brief rationale:

- Document processes to monitor for and detect test irregularities, evidence of cheating, and evidence of inadvertent item release detectable at the CAI level, and processes for communicating these in a timely fashion to all participating states.

- Evidence regarding processes for remediation following any test security incidents involving loss of security over specific test items during item development, standard-setting or other test development activities that are under the control of CAI.

Critical Element 2.6 – Systems for Protecting Data Integrity and Privacy

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or 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: • (CE 2.6.1) To protect the integrity of its test-related data in test administration, scoring, storage and use of results;	CE 2.6.1: ● *CS 005 – V5_Shared Technical Report.pdf* ● *CS 002.E – Appendix E. Sample Data Review Training Materials.pdf* ● *CS 003.B – Appendix B. Standard-Setting Training Slides.pdf*	CE 2.6.1: While *CS 006 – V6* briefly describes the types of scores students might receive, no evidence is found with regard to CAI having created systems or guidance for states in the protection of student data. The websites and software systems used by CAI during development utilized features to protect authorized data access through password-protected logins during item development and review, test delivery, and score reporting. (CS 005 – V5, p 5). It is not clear where the data are stored and how security is maintained within this storage setup. In addition, it is not clear how hand scoring is done for the paper-pencil version of the test, so there is no evidence regarding policies and procedures to protect the integrity of test-related data and materials during scoring. **States need to provide evidence for the protection of their test-related data in test administration, scoring, storage and use of results.**
• (CE 2.6.2) To secure student-level assessment data and protect student privacy and confidentiality, including guidelines for districts and schools;	CE 2.6.2: ● *CS 005 – V5_Shared Technical Report.pdf* ○ Section 3.1-Provides FERPA compliance (p. 5; pdf p.7)	CE 2.6.2: CAI provides guidance to states regarding practices that are forbidden as part of FERPA legislation, including practices regarding login information, passwords, sending personal student information via email, and State Student ID numbers.

• (CE 2.6.3) 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.	CE2.6.3: ● *CS 005 – V5_Shared Technical Report.pdf* o Section 3.1-Provides FERPA compliance (p. 5; pdf p. 7)	**States need to provide evidence that they implemented the guidelines recommended by CAI as part of FERPA requirements.** No evidence is found regarding guidance from CAI to ensure procedures are in place for compliance when human assistance is needed during testing, such as a human reader or human assistance with other accommodations. No evidence is found regarding firewalls or other provisions for ensuring that each states' data is separated from other states' data. CE 2.6.3: CAI provides the caution to its participating states that test materials and score reports should not be exposed to identify student names with test scores, except by authorized individuals with a need to know. **Individual states must provide evidence regarding how they define minimum cell size for reporting.**

Section 2.6 Summary Statement

___ No additional evidence is required or

_X_ The following additional evidence is needed/provide brief rationale:

- Evidence regarding where the data are stored and how security is maintained within this storage setup.
- Evidence of guidelines to states for protecting student information in all testing/scoring situations (e.g., human assistance during testing, hand scoring of paper/pencil tests).
- Evidence regarding firewalls or other provisions for ensuring that each states' data is separated from that of other states.

SECTION 3: TECHNICAL QUALITY – VALIDITY

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

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State has documented adequate overall validity evidence for its assessments consistent with nationally recognized professional and technical testing standards. The State's validity evidence includes evidence that: **<u>The State's academic assessments</u>** measure the knowledge and skills specified in the State's academic content standards, including: • (CE 3.1.1) 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), balance of content, and cognitive complexity;	CE 3.1.1: • *CS 002 – V2_Shared Technical Report.pdf* o Introduction – Overviews the development of the Shared Science Item Bank and discusses the role of the item bank development process in ensuring valid state assessments (p. 1; pdf p. 4) o Section 1.1 – Discusses the claims the Shared Science Item Bank and resulting state-level tests would be designed to support (p. 2; pdf p. 5) o Section 2.1 (Exhibit A) – Shows the development steps supporting the validity of claims (p. 4; pdf p. 7) o Section 2.2 – Describes the participation of states in developing item specifications aligned to three-dimensional performance expectations (pp. 4-7; pdf pp. 7-10) o Section 3 – Describes the composition of the resulting item bank and demonstrates how it is suited to assessments based on a three-dimensional conceptualization of science understanding (pp. 18-25; pdf pp. 21-28) o Section 3.2 – Demonstrates CAI's overall item development plan (p. 25; pdf p. 29)	CE 3.1.1: CAI provides evidence regarding the item development process designed to build items from item specifications reflecting the three-dimensionality of the Next Generation Science Standards (NGSS) and the framework of all the participating states' standards. The item specifications (*CS 002.K – Appendix K. Item Specifications – Grade 3-High School)* are well detailed, including the three dimensions to be assessed, content limits and grade-level vocabulary, as well as task demands. The item specifications taken as a whole provide evidence of the range and balance of content covered and the levels of cognitive complexity to be assessed. The most compelling evidence for standards alignment was found in *CS 004.C – Appendix C. Alignment Report.* This document provides evidence of alignment to the standards in respect to the three dimensions, cognitive complexity, and item/cluster distribution. Evidence of where adjustments and improvements to alignment are needed to meet the full depth and breadth of the standards was also included.

	○ Section 4 – Provides recommendations and guidelines for states to develop test blueprints and describes the test designs supported by CAI's administration platforms (used by the majority of partnering states) (pp. 25-27; pdf pp. 29-31) ● *CS 002.K – Appendix K. Item Specifications – Grade 3-High School.pdf* Includes detailed item specifications by Performance Expectation for each grade level ● *CS 003 – V3_Shared Technical Report.pdf* • Section 5 – Describes the content-based standard-setting methodology (pp. 5-20; pdf pp. 7-22) *CS 004.C – Appendix C. Alignment Report.pdf* ○ Executive Summary – Describes the research questions for the alignment study, including breadth and depth of item bank coverage (pp. 2-3; pdf pp.7-8) ○ Table 4: Consensus Alignment Criteria – Describes the criteria on which alignment was evaluated (pp. 30-31; pdf pp. 35-36) ○ Alignment Criteria – Describes the criteria by which results are reported (pp. 32-37; pdf pp. 37-42) ○ Summary Findings – Provides a summary of analyses and results (pp. 38-50; pdf pp. 43-55) ○ Conclusions – Provides a response to each identified research question, as well as recommendations for addressing areas of concern (pp. 53-60; pdf pp. 58-65)	In terms of balance of content, it is not clear why there are so few items for the Engineering, Technology, and Applications of Science discipline (*CS 002 – v2*, pdf pg. 25, 27), and whether this is in balance with the proportion of Engineering, Technology and Applications standards in the NGSS. The various item types used in the assessment support the development of items representing a full hierarchy of cognitive complexity (*CE 004.C*, p. 11-12). It is unclear, however, the extent to which the different item types were able to capture all levels of cognitive complexity, and to what extent the SEPs and CCCs were used to contribute to cognitive complexity at all the grade levels. An independent alignment study was conducted (*CS 004.C – Appendix C. Alignment Report, p. 53))* that provides evidence showing that the Shared Science Assessment Item Bank has the capacity to generate aligned test forms, pending **state-specific evidence of alignment studies based on their specific versions of the assessment.** One alignment weakness identified in the alignment study (p 53) was in the high school item pool, indicating the need for an additional six items in order to target unrepresented Performance Expectations (PEs). Within document *CS 002.H – Appendix H. Science Item Bank,* PEs are listed for which there were no testable items. This suggests the existences of gaps in assessing student knowledge and skill in that standard area. PEs that indicate a zero (0) include: HS-ESS2-5, HS-ESS3-4, HS-PS2-6, HS-PS4-3, and HS-PS4-4. There are also PEs that are addressed by only a single test item. Evidence is needed to document whether CAI has developed the additional items needed to meet this deficiency. The alignment study (p. 59) noted many instances of editorial errors throughout the items as well as a variety of

• (CE 3.1.2) Documentation that the assessments address the depth and breadth of the content standards;	CE 3.1.2: • *CS 002 – V2_Shared Technical Report.pdf* o Section 3 – Describes the content of the Shared Science Assessment Item Bank (pp. 18-25; pdf pp. 21-28) o Section 4.2 – Describes the shared characteristics of all states' blueprints that ensure depth and breadth of coverage (p. 26; pdf p. 30): ▪ The number of performance expectations that can be tested only once ▪ The general rules about the number of item clusters or stand-alone items that can be sampled from the same Disciplinary Core Idea ▪ Any specific constraints that might be imposed by the state o Section 5 – Describes the process for generating item exposure reports for the purpose of monitoring and planning future item development (pp. 27-28; pdf pp. 31-32)	other questionable aspects of specific items in the bank. It is not clear if and how CAI has addressed these various concerns. CE 3.1.2: CAI provided evidence of processes used in item development to address the depth and breadth of the content standards, including the item developer training, the detailed item specifications, and the alignment study designed to check that the set of items used in test simulations met the depth and breadth of the standards. In addition, CAI provided evidence that the processes used during item development in fact resulted in the desired number of items in the shared bank in terms of grade band, type of item (cluster or stand-alone), science discipline, and disciplinary core idea (DCI). (*CS 002 – v2*, pg. 21-24). What is not provided is evidence indicating: (a) Number of items by SEPs and CCCs. (b) Number of items by item type (e.g., choice, text entry, table, edit text, grid). Such information would provide helpful evidence regarding the extent to which the assessment fully addresses the depth and breadth of the standards. While CAI describes a simulation process for generating the optimal item exposure rates for each assessment and to evaluate whether individual tests adhere to each state's blueprint (*CS 002-v2*, p. 27), it is not clear what process will be followed if a state's assessment is not a good match for its blueprint or when and how future items will be developed to address any needs that surface. **States will need to document evidence of the depth and breadth of standards coverage for their individual state's assessment via their test blueprints, their choice of test format, item exposure rates, and alignment**

• (CE 3.1.3) If the State has adopted alternate academic achievement standards and administers alternate assessments aligned with those standards, the assessments show adequate alignment to the State's academic content standards for the grade in which the student is enrolled 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.	CE 3.1.3: Not applicable	**studies conducted specifically on their version of the assessment.** CE 3.1.3: Not applicable

Section 3.1 Summary Statement

___ No additional evidence is required or

_X_ The following additional evidence is needed/provide brief rationale:

- Evidence regarding follow-up to the finding in the alignment study indicating the need for additional high school items to target underrepresented Performance Expectations (PEs).
- Evidence that the large number of editorial issues discovered by the alignment study were addressed.
- Evidence that the assessment includes appropriate coverage of the SEPs and CCCs in terms of reflecting the depth and breadth of the standards.

Critical Element 3.2 – Validity Based on Cognitive Processes

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or 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.	• *CS 001 – V1_Shared Technical Report.pdf* o Section 1.2– Describes common elements of participating states' adopted standards (p. 2; pdf p. 7) o Section 2.3 – Describes how embedded accessibility features in the administration system underly an assessment that taps the intended cognitive processes without the influence of construct irrelevant impediments (pp. 5-6; pdf pp. 10-11) o Section 3 – Summarizes content review, rubric validation, and field test data review procedures during development (p. 7; pdf p. 12) o Section 4 – Summarizes field test data review (pp.19-23; pdf pp. 24-28) o Section 7 – Summarizes how item performance is monitored during operational administrations (p. 41; pdf p. 46) • *CS 001.A – Appendix A. Classical Statistics for Science Items.xlsx* o Provides field test data at the assertion and item levels aggregated by state (i.e., p-value, point biserial correlation and DIF) • *CS 001.B – Appendix B. Science Calibration Item and Group Parameters.xlsx* o Provides operational item parameters for 2018 and 2019 • *CS 002 − V2_Shared Technical Report.pdf* o Section 2.2 – Establishes the intended cognitive processes that guide item writing (pp. 4-7; pdf pp. 7-10) o Section 2.4 – Describes internal iterative review process to ensure items measure the	The independent alignment study (*CS 004 – V4)* documented analyses that found acceptable alignment between the item bank and grade-level expectations across the standards, matching for the most part the three dimensions of the standards that encompass the cognitive processes appropriate at each grade level. In the Cognitive Lab studies (*CS 004 A* and *B*), students reported that the tests seemed appropriate for their grade level. **States need to provide evidence that their state-specific assessments tap the intended cognitive processes appropriate for each grade level**. *CS 004.C – Appendix C. Alignment Report* notes that the majority of questions for grade 5 and high school represent complexity Level 2. The middle school assessment questions range between Levels 2 and 3. Level 4 questions are barely represented in the item bank. Although the report indicates that these results were not considered a threat to test alignment, **States using items and clusters from this assessment bank need to be aware of this limitation to ensure they are meeting the depth and breadth of their science standards as well as content complexity**. **Test developers and States should also consider this information as they build more items and clusters to ensure all levels of cognitive processing are covered at all grade levels.**

	intended cognitive process as established by the item specifications (pp. 7-10; pdf pp. 10-13) o Section 2.5 – Describes multi-state involvement in content review process (pp. 10-13; pdf pp. 13-17) o Section 2.7 – Describes data review and rubric validation processes (pp. 14-18; pdf pp. 17-21) • *CS 002.K – Appendix K. Item Specifications – Grade 3-High School.pdf* o Describes item specifications by grade level for each Performance Expectation. Identifies content and content limits and task demands to provide specific guidance to item writers • *CS 002.L – Appendix L. Style Guide for Science Items.pdf* o Style specifications ensure items tap intended cognitive processes • *CS 004 – V4_Shared Technical Report Volume 4.pdf* o Section 1.2 – Validity – Describes the first source of validity as the relationship between test content and the intended test construct, including the required cognitive skills (p. 3; pdf p. 6) o Section 4.2 – Independent Alignment Study – Describes the independent alignment study at the item bank level (p. 9; pdf p. 12) o Section 6.1 – Cognitive Laboratory Studies - Describes the cognitive lab studies (pp. 19-20; pdf pp. 22-23) • *CS 004.A − Appendix A. Science Clusters Cognitive Lab Report.pdf* • *CS 004.B − Appendix B. Braille Cognitive Lab Report.pdf* • *CS 004.C – Appendix C. Alignment Report.pdf* o Conclusions – Research Question 2: "What	

	Category of Engagement (cognitive complexity) is required for successful completion of each interaction with a stand-alone item or item cluster and how does this compare with the Category of Engagement assigned to the corresponding PE?" – Addresses whether items in the Shared Science Assessment Item Bank address the intended cognitive processes. Results state, "Out of the 389 items/clusters included in the analysis, only five items (1%) were flagged for revision or removal with the primary issue identified related to the Category of Engagement." (p. 55; pdf p. 60)	

Section 3.2 Summary Statement

___ No additional evidence is required or

_X_ The following additional evidence is needed/provide brief rationale:

- Evidence regarding follow-up to the finding in the alignment study indicating the need for additional high school items to target underrepresented PEs, as noted in CE 3.1
- Evidence that the item pool is sufficiently deep to tap the highest level of intended cognitive processes (e.g., Cognitive Level 4) and/or plans to address this need in future item development.

Critical Element 3.3 – Validity Based on Internal Structure

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or 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**.	• *CS 001 – V1_Shared Technical Report.pdf* o Section 4.4 – Includes formulas and flagging criteria for detecting DIF in field-tested items (pp. 21-22; pdf pp. 26-27) • *CS 004 – V4_Shared Technical Report.pdf* o Section 5.1 – Discusses an examination of cluster effects as an indicator of dimensionality (pp. 9-18; pdf p. 12-21) o Section 5.2 – Discusses recommendations for examination of convergent and divergent validity at the test level (pp. 18-19; pdf pp. 21-22) o Section 5.1.2 – Shows cluster variance from IRT calibrations for the entire bank (p. 10; pdf p. 13) o Section 5.1.3-5.1.4 – Study used Utah's fixed form approach in 2018 to evaluate the internal factor structure (p. 10-18; pdf pp. 13-21) • *CS 003 – V3_Shared Technical Report.pdf* o Section 4.2 – Shows sample item cluster and defensible scoring assertions for an assessment based on three-dimensional science standards (p. 3; pdf p. 6) o Section 5.1 – Discusses psychometrically sound procedures for determining performance standards based on the structure of three-dimensional science standards. (pp. 3-4; pdf pp. 8-9)	CAI conducted a series of validity studies (*CS 004 – v4*, pgs. 9-19) examining internal-external structure match, including (1) simulation studies confirming cluster effects, and (2) a factor analysis of Utah's fixed-form test. These analyses generally supported the validity of the test results being consistent with the sub-domains, except for two models used in the factor analysis study (p. 17). The misfit for these two models was caused by one item in all the grade 6 forms examined, and when that item was removed from the analysis, the measures of fit improved. It is unclear if CAI removed this item from the item bank. A significant question is what CAI considers the sub-domain structure of the standards in reference to this assessment. In one case, it appears to be based on PEs which would include all three dimensions, while in other cases, the disciplines (life, physical, Earth and space) are used as the sub-domain reference. Since studies were conducted with both of these sub-domain structures, the evidence appears to be more exploratory, aimed at identifying which sub-domains might show the best validity, rather than CAI having built the assessment with a pre-determined sub-domain structure of the standards in mind. The sub-domain structure that would be most appropriate would be the three dimensions of the academic content standards (DCIs, SEPs, and CCCs). **Because participating states have state-specific standards and blueprints, each state must provide evidence in support of validity based on internal structure of the assessment based on its unique test**

		make-up and scoring and reporting structures, and the structure of the content standards in that state.

Section 3.3 Summary Statement

___ No additional evidence is required or

_X_ The following additional evidence is needed/provide brief rationale:

- Clarity regarding the actual sub-domain structure of the standards against which the internal structure of the assessment is being validated, and a rationale why no analysis was conducted to look for adequate validity evidence that the scoring and reporting structures of the assessment are consistent with the three-dimensional structure of the academic content standards that the assessment is designed to measure.

Critical Element 3.4 – Validity Based on Relations to Other Variables

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State has documented adequate validity evidence that the State's assessment scores are related as expected with other variables.	• *CS 004 – V4_Shared Technical Report.pdf* o Section 1.2 – Describes evidence to support validity based on relations to other variables (p. 4; pdf p.7) o Section 5.2 – Describes evidence for convergent and discriminant analysis (pp. 18-19; pdf pg. 21-22)	Since CAI conducted the validity studies with one state's data, it is up to each p**articipating state to submit evidence that the scores resulting from their unique assessments are related as expected with other variables.** Of the analyses that were conducted by CAI, some questions arose with regard to: • Specific correlations with other science assessments. • Specific correlations between sub-scores and with ELA and math. • Clarity regarding whether students in the Cognitive Labs were reporting their performance strategies in terms of answering the substantive content part of the assessment or their strategies in interacting with the technological features of the assessment.

Section 3.4 Summary Statement

X_ No additional evidence is required or
___ The following additional evidence is needed/provide brief rationale:

SECTION 4: TECHNICAL QUALITY – OTHER

Critical Element 4.1 – Reliability

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or 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 consistent with nationally recognized professional and technical testing standards. If the State's assessments are implemented in multiple States, measures of reliability for the assessment overall and each student group consistent with nationally recognized professional and technical testing standards, including: • (CE 4.1.1) Test reliability of the State's assessments estimated for its student population;	CE 4.1.1: • *CS004 – V4_Shared Technical Report.pdf* o Section 1.1 – Provides an overview of reliability determinations including SEM (classical test theory) and TIF (item response theory) (pp. 2-3; pdf pp. 5-6) o Section 3 – Describes the computation reliability model used (specific to LOFT test designs) (pp. 5-8; pdf pp. 8-11) • *CS001 – V1_Shared Technical Report.pdf* o Section 5.1 – Describes the structure of the IRT model. (pp. 23-26; pdf pp. 28-31) o Section 6.4 – Describes the context for interpretation of the reliability coefficients (i.e., the LOFT test design) (p. 38; pdf p. 43) • *CS 001.A – Appendix A: Classical Statistics for Science Items*	CAI determined reliability for each state following an operational administration in the state. **States will need to provide the reliability evidence in their individual state submissions.** CE 4.1.1: CAI ran extensive statistical analyses at both the item and the aggregate level. This is especially important since students are not given a fixed form of the test, so it is important to know how each item behaves as well as the test as a whole. The analyses used are consistent with nationally recognized professional and technical testing standards. Because the items in the item bank are not unidimensional, and therefore have interdependence of assertions, CAI ran analyses both using traditional methods and also analyses in which the local dependencies are ignored. Across states and grades, the overall reliability ranged from 0.84 to 0.89. When the local dependencies among assertions related to the same item were ignored, the marginal reliability increased to 0.90 and upward. (*CS 004 – v4, pg. 6*). *CS 001.A – Appendix A* provides classical statistics for all items in the bank, by state and by student sub-groups (sex, race, and Special Education).
• (CE 4.1.2) Overall and conditional standard error of measurement of the State's assessments, including any domain or component sub-tests, as applicable;	CE 4.1.2: • *CS 001 – V1_Shared Technical Report.pdf* o Section 6.4 – Describes the computation method of CSEM (p. 38; pdf p. 43) • *CS004 – V4_Shared Technical Report.pdf*	CE 4.1.2: CAI describes the process used for computing the standard error of measurement. **Because conditional standard errors are test and population specific, they are reported in each state's technical report.**

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
• (CE 4.1.3) Consistency and accuracy of estimates in categorical classification decisions for the cut scores, achievement levels or proficiency levels based on the assessment results; • CE 4.1.4) For computer-adaptive tests, evidence that the assessments produce test forms with adequately precise estimates of a **student's academic achievement**.	○ Section 3.3 – Refers to the mean CSEM at each performance level (p. 8; pdf p. 11) CE 4.1.3: • *CS004 - V4_Shared Technical Report.pdf* ○ Section 3.2.2 – Describes the classification consistency index (p. 8; pdf p. 11) ○ Section 3.3 – Provides a description of the mean CSEM at each performance level (p. 8; pdf p. 11) CE 4.1.4: ● *CS002.J – Appendix J. Adaptive Algorithm Design.pdf* ○ Section 1.3.3 – Provides description of the item selection for the LOFT design (p. 7)	CE 4.1.3: CAI describes the computational methods used for calculating both the classification accuracy and the classification consistency for each state's version of the test, (*CS 004 -V4, pg. 6-8*), as well as the mean CSEM by grade level for each state. **States will need to report the specific statistics in terms of classification accuracy and classification consistency for the performance levels as defined for their state science assessment.** CE 4.1.4: CAI describes in detail (*CS002.J – Appendix J. Adaptive Algorithm Design)* the algorithm used for generating items for each student, designed to maximize alignment to the blueprint as well as to estimate the student's proficiency as closely as possible. **States will need to report evidence that when the algorithm is applied on their state assessment, acceptable levels of precision are found for estimating individual student achievement.**
Section 4.1 Summary Statement		
_X_ No additional evidence is required or ___ The following additional evidence is needed/provide brief rationale:		

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 4.2 – Fairness and Accessibility

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
For all State academic assessments, assessments should be developed, to the extent practicable, using the principles of universal design for learning (UDL) (see definition[1]). **<u>For academic content assessments</u>**, the State has taken reasonable and appropriate steps to ensure that its assessments are accessible to all students and fair across student groups in their design, development and analysis.	• *CS 001 – V1_ Shared Technical Report.pdf* o Section 1.4 – Describes special test formats and versions available for students unable to participate in online administration, including braille, American Sign Language (ASL), closed captioning, and large print (pp. 3-4; pdf pp. 8-9) o Section 3.1 – Describes the field testing process (pp. 7-18; pdf pp. 12-23) o Section 4 – Describes data review processes and flagging criteria, including DIF (pp. 19-20; pdf pp. 24-25) • *CS 002 − V2_Shared Technical Report.pdf* o Section 1.2 – Identifies underlying principles for item development, stating, "Items were written with the goal that virtually every item would be accessible to all students, either by itself or in conjunction with accessibility tools, such as text-to-speech, translations, or assistive technologies. This goal is supported by delivery of the items on CAI's test delivery system, which has received Web Content Accessibility Guidelines (WCAG) 2.0 AA certification…" (p. 2; pdf p. 5) o Section 2.1 – Exhibit A identifies accessibility considerations for each step of the item development process (p. 4; pdf p. 7) o Section 2.3 – Provides selection and training of item writers (p. 7; pdf p. 10)	Universal Design for Learning (UDL) principles are evident as part of the training conducted for item writers. It is also evident in the range of item types available for use in the test, ranging from selected-response items to grid in items, to short answer and more extended responses. In other words, across the assessment, students have a variety of ways in which they might be asked to demonstrate their knowledge and skills. As part of the review process during item development, CAI included reviews for language accessibility, bias, and sensitivity. Differential Item Functioning (DIF) analyses were conducted to identify and address differences in student performance by sub-group (e.g., sex, ethnicity) that may be due to bias in items that would make it difficult for certain students to access the item in the way it was intended. CAI included embedded and non-embedded universal tools for all students, designated supports for students whose teachers or other staff thought would benefit them, and accommodations specific to students with IEPs or in EL programs. The range of supports for students reflect the common supports and accommodations typically allowed on state assessments so that students who have taken other state assessments should be familiar with the various supports provided.

[1] see page 28 of "*A State's Guide to the U.S. Department of Education's Assessment Peer Review Process"*, September 24, 2018 available at: www.ed.gov/admins/lead/account/saa.html

	○ Section 2.4 – Describes internal review process (pp. 7-10; pdf pp. 10-13) ○ Section 2.5 – Includes state review process (pp. 10-13; pdf pp. 13-16) ○ Section 2.7.2 – Describes field test data review process including DIF (pp. 15-18; pdf pp. 18-21) ● *CS 002.A – Appendix A. Item Writing Training Materials.pdf* ○ Provides item writer training materials ● *CS 002.B – Appendix B. Item Review Checklist.pdf* ○ Provides the review checklist developed by CAI ● *CS 002.C – Appendix C. Content Advisory Committee Participant Details.pdf* ○ Provides the state review participants ● *CS 002.D – Appendix D. Fairness Committee Participant Details.pdf* ○ Provides the fairness review participants ● *CS 002.E – Appendix E. Sample Data Review Training Materials.pdf* ● *CS 002.F – Appendix F. Data Review Committee Participant Details.pdf* ● *CS 002.L – Appendix L. Style Guide for Science Items* ○ Establishes consistent style to ensure clarity of language and accessibility within test items ● *CS 004 – V4_Shared Technical Report.pdf* ○ Section 6.1 – Summarizes the purposes of the cognitive laboratories conducted during item development (pp. 19-20; pdf pp 22-23) ● *CS 004.A – Appendix A – Science Clusters Cognitive Lab Report.pdf* ○ Provides detailed results of cognitive laboratory study ● *CS 004.B – Appendix B – Braille Cognitive Lab Report.pdf*	*CS 005* indicates that CAI created a practice test to ensure that students become familiar with and can use all the tools necessary to engage with the testing platform. **States will need to document that requirements are in place to ensure that all students are given opportunities to take the practice test to ensure they are comfortable accessing and using the online formatting, tools, and settings.** The Cognitive Lab studies in general provided evidence that students felt that the items were accessible to them both in terms of using the functionality of the online engine and in terms of grade-level appropriateness. The Cognitive Lab Braille study did raise a concern: • In the Braille Cognitive Lab study, students reported a number of formatting and technology-related issues that negatively impacted their ability to optimally demonstrate their knowledge and skills. Recommendations for addressing these issues are listed in *CS 004.B – Appendix B – Braille Cognitive Lab Report*. pg. B- 8,9. It is unclear if CAI made adjustments to the test and/or to the test engine to address the deficiencies in the Braille version of the assessment. If those modifications have not been made, the validity of the results of the assessment for visually-impaired students would be suspect. With regard to the focus on vocabulary in the item specifications, it is unclear if students are assessed based on their understanding of science content or on their knowledge of vocabulary. While *CS 005* lists the use of a word-to-word bilingual dictionary as an accommodation for ELs, no evidence was found for the inclusion of an embedded dictionary in the assessment to provide meanings of words that are not expected at that grade level or as part of the standards. Otherwise, vocabulary used in the assessment could be discriminatory for some students, especially ELs and students with disabilities.

	○ Provides detailed results of braille cognitive laboratory studies ● *CS 005 – V5_Shared Technical Report.pdf* ○ Section 1.2 – Provides testing options (p. 1-2; pdf pp. 3-4) ○ Section 5 – provides accommodations details (pp. 7-13; pdf pp. 9-15)	It would be very informative in terms of ensuring fairness and accessibility for all students for CAI to conduct studies (e.g., Cognitive Labs) in the future to examine accessibility to the assessment for EL students and for students with various disabilities.

Section 4.2 Summary Statement

___ No additional evidence is required or

_X_ The following additional evidence is needed/provide brief rationale:

- Evidence that the deficiencies in the Braille version of the assessment that hindered the visually-impaired students' accessibility to the assessment have been rectified.
- Evidence that vocabulary used throughout the assessment is accessible and appropriate for all students, including attention to vocabulary being correctly aligned to the standards by grade level, and consideration of the inclusion of an embedded dictionary for any uncommon or figurative vocabulary used on the assessment that is not being assessed as part of the standards but may be unknown to ELs or other students.
- Evidence of planning future Cognitive Labs to assess fairness and accessibility issues for ELs and students with disabilities (other than the visually-impaired).

Critical Element 4.3 – Full Performance Continuum

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence

The State has ensured that each assessment provides an adequately precise estimate of student performance across the full performance continuum for **academic assessments**, including performance for high- and low-achieving students.	• *CS 001 – V1_Shared Technical Report.pdf*
	  ○ Section 5.1 – Describes the IRT model for item calibration and computing scores (pp. 23-26; pdf pp. 28-31)
	  ○ Sections 6.1 and 6.2 – Discuss scaling (pp. 35-38; pdf pp. 40-43)
	  ○ Section 6.3 – Describes extreme case handling (specific methods and HOT and LOT values are reported in state-specific technical reports) (p. 38; pdf p. 43)
	  ○ Section 6.4 – Describes how the conditional standard error of measurement (CSEM) is calculated (p. 38; pdf p. 43)
	  ○ Section 6.5 – Describes calculation of reported scale scores (p. 38-39; pdf pp. 43-44)
	• *CS 002 – V2_ Shared Technical Report.pdf*
	  ○ Section 2.2 – Defines content limits for item development: "…delineates the specific content that the PE measures and the parameters in which items must be developed to assess the PE accurately, including the lower and upper complexity limits of items." (p. 5; pdf p. 8)
	  ○ Section 5 – Describes simulation summaries (pp. 27-28; pdf pp. 31-32)
	• *CS 002.K – Appendix K. Item Specifications – Grade 3-High School.pdf*
	  ○ Clearly defines content limits for each Performance Expectation, ensuring development of items that are appropriate for students across the entire performance continuum
	• *CS 003 – V3_Shared Technical Report.pdf*
		CAI has provided industry-accepted procedures by which **each state should conduct their test-level analyses:**
		• *Shared Technical Report Volume 1* describes scaling methods, including determination of extreme case scores, and conditional standard errors of measurement (CSEM).
		• *Shared Technical Report Volume 3* describes how to determine the scale score cut defining each proficiency level.
		• *Shared Technical Report Volume 4* Section 3 describes how to compute reliability and the conditional standard errors of measurement. The scores are expected to be most precise near the proficient level cuts and somewhat less so for students at the two ends of the continuum.
		
		State-level technical reports will need to document evidence of adequately precise estimates of student performance across the full performance continuum, including for high- and low-achieving students for the final operational assessment used in their state, including such evidence as:
		• score frequency distributions
		• test-level alignment studies
		• operational SEM and CSEM computations
		• Test Information Functions
		• indices of classification accuracy and consistency relative to the state's specific blueprint and scoring structures.
		
		The appendices to *CS001* provide empirical evidence for a range of item difficulties that would be appropriate for high- and low-achieving students.
		
		Good information is provided on the CAT, demonstrating

	○ Section 5.7.1 – Describes calculation of scale score cuts for proficiency (pp. 15-16; pdf pp. 18-19) ● *CS 003.A – Appendix A. Development of Range Performance-Level Descriptors.pdf* ○ Describes the process for developing PLDs across the performance continuum ● *CS 004 – V4_Shared Technical Report.pdf* ○ Section 3 – Provides the reliability, CSEM, classification accuracy and consistency computations (pp. 5-8; pdf pp. 8-11) ○ Section 4.2 – Summarizes results of the independent alignment study of the alignment bank (p. 9; pdf p. 12) ● *CS 004.C – Appendix C. Alignment Report.pdf* ○ Research Questions 1 and 2 and corresponding analyses address the potential for items to address student performance across the entire performance continuum, as defined by the intended cognitive complexity of the standards (pp. 54-55, pdf pp. 59-60) • *CS 001.A - Appendix A. Classical Statistics for Science Items* • *CS 001.B - Appendix B. Science Calibration Item and Group Parameters*	the initial item selection on the full continuum and then narrowing into the location on the scale where student performance indicators they fall. Providing the CSEMs at the tails and cut points for a few sample forms would be helpful, particularly for the LOFT form. **State-level data on the SEMs are needed to completely evaluate this CE. The state results need to show adequate precision in order for this CE to be fully met.**

Section 4.3 Summary Statement

_X_ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:

Critical Element 4.4 – Scoring

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State has established and documented standardized scoring procedures and protocols for its assessments that are designed to produce reliable and meaningful results, facilitate valid score interpretations, and report assessment results in terms of the State's **academic achievement standards**.	• *CS 001 – V1_Shared Technical Report.pdf* o Section 5.1, 5.2 and 5.3 – Describes the IRT model, the item calibration process, and linking procedures used to correlate the 2018 to the 2019 scale (pp. 23-35; pdf pp. 28-40) o Section 6.1, 6.2 and 6.3 – Describe scale score calculation (pp. 35-38; pdf pp. 40-43) o Section 6.4 – Describes calculation of Standard Errors of Estimate (p. 38; pdf p. 43) o Section 6.5 – Describes the process used to establish the student-level scale in each participating state (pp. 38-39; pdf pp. 43-44) o Section 6.6 – Describes rules for calculating performance levels (p. 39; pdf p. 44) o Section – 6.7 – Describes possible reporting levels based on scale scores (pp. 40-41; pdf pp. 45-56) o Section 7.1.4 – Describes score production via CAI's scoring engine and affirms that a second scoring verification system is used to verify that all test scores match with 100% agreement in all tested grades. (p. 42; pdf p. 47) ● *CS002 – V2_Shared Technical Report.pdf* o Section 2.7.1 – Describes the rubric validation process used for scoring accuracy (p. 14-15; pdf p.17-18) ● *CS 003 – V3_Shared Technical Report.pdf* o Section 2 – Provides the purpose and overview for the common standard-setting methodology (p. 1; pdf p. 4) o Section 5 – Provides the standard-setting process (pp. 3-5; pdf pp. 7-8) ● *CS006 – V6_Shared Technical Report.pdf*	**Scoring for each participating state is conducted under the terms of the specific contract between the state and CAI.** **States will need to provide documentation regarding rules for invalidating test results when necessary.** It is unclear how scoring was designed and implemented for open-ended items. No samples of open-ended items were given, nor an explanation about how the machine was programmed for scoring open-ended items. Evidence was not found with regard to using human scoring to develop the rubrics, and no comparability study was conducted to determine the reliability of machine scoring in comparison to human scoring. It is also unclear how the scores will be reported to parents in terms of expectations for students. It is also unclear how the scoring procedures will be used to support valid score interpretations, especially in terms of the three dimensions of the standards. It is unclear how scoring will be conducted for tests that are taken using the paper version of the assessment, and how those scorers will be trained.

	○ Section 1.2 – Describes state reporting systems overall (p. 2; pdf p. 4) ○ Section 2 – Describes interpretation of reported scores (p. 2; pdf p. 4)	

Section 4.4 Summary Statement

___ No additional evidence is required or

_X_ The following additional evidence is needed/provide brief rationale:

- Evidence regarding the development and implementation of scoring for open-ended items to ensure reliable and meaningful results.
- Evidence that clarifies how scores will be used to facilitate valid score interpretations and report assessment results in terms of the State's academic achievement standards.

Critical Element 4.5 – Multiple Assessment Forms

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
If the State administers multiple forms of **academic assessments** 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.	• *CS 001 – V1_Shared Technical Report.pdf* o Section 3.2 – Describes test assembly and item selection within the LOFT test design (pp. 18-19; pdf pp. 23-24) o Section 5.2 and 5.3 – Describes item calibration and linking of the scale to subsequent administrations (pp. 26-35; pdf pp. 31-40) • *CS 002 − V2_Shared Technical Report.pdf* o Section 3.2 – Describes the strategy for replenishing the item bank based on identified gaps in coverage and item exposure (ensures availability of items to fill blueprint constraints across multiple administrations) (p. 25; pdf p. 29) o Section 4.1 – Describes test design (LOFT and CAT capability of the administration platform) and functioning of the LOFT algorithm (p. 25-26; pdf pp. 29-30) o Section 5 – Describes simulation studies and their results (pp. 27-28; pdf pp. 31-32)	Each state uses a LOFT (linear on the fly) design to generate different forms of the assessment to each student. The decision-model for which items are pushed out to each student are based on an algorithm designed to ensure comparability of standards coverage and test blueprint match across forms. CAI ran simulation studies to check that the algorithm used in the LOFT design produced online tests that were comparable from student to student and from administration to administration. **The simulations were carried out for each state, so each state will need to report the findings of the simulation studies appropriate for them.** (*CS 002 – V2*, p. 27). The explanation of comparability within and across LOFT and CAT designs seems reasonable. It would be helpful if CAI conducted a simulation across the item bank that is not state-specific to provide evidence of the validity of the process being used. Item parameter estimates obtained from the 2018 student responses were highly correlated with the item parameters obtained from the 2019 student responses (*CS 001-V1*, p 34). The statistics for these analyses are reported. CAI describes (*CS 002-V2*, p 25) types of items that will need to be replaced, and mentions that CAI and participant states continue to develop items to "replenish and grow" the item bank, but it is not clear how often this is done, who does the continued item development, or what kind of training and review is used in the continued item development process. As the item replenishment process ensues, it is unclear how comparability across forms will continue to be validated.

Section 4.5 Summary Statement

___ No additional evidence is required or

_X_ The following additional evidence is needed/provide brief rationale:

- Clarity regarding details for continued item development (e.g., timeline, training, review) and how comparability across forms and across years will continue to be validated.

Critical Element 4.6 – Multiple Versions of an Assessment

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
If the State administers any of its assessments in multiple versions within a subject area (e.g., online versus paper-based delivery; **or a native language version of the academic content assessment**), grade level, or school year, the State: • (CE 4.6.1) Followed a design and development process to support comparable interpretations of results for students tested across the versions of the assessments;	CE 4.6.1: • *CS 001 – V1_Shared Technical Report.pdf* o Section 1.4 – Describes test formats and special versions (pp. 3-4; pdf pp. 8-9) o Section 2.3 – Describes universal features, designated supports and accommodations (pp. 5-6; pdf pp. 10-11) • *CS 002 – V2_Shared Technical Report.pdf* o Section 2.2 – Provides a list of claim structures and principles guiding development (pp. 4-7; pdf pp. 8-10) o Section 2.5.4 – Describes markup of items in the Shared Science Assessment Item Bank for Spanish translation and braille for a subset of items (p. 13; pdf p. 16) o Section 4.4 – Describes the common features across participating states for paper-pencil form construction (pp. 26-27; pdf pp. 30-31) • *CS 004 – V4_ Shared Technical Report.pdf* o Section 6.1 – Overviews a cognitive laboratory study of the braille assessment forms supporting similar performance of items for students using braille and students responding to the items in the online environment (pp. 19-20; pdf pp. 22-23) • *CS 004.B – Braille Cognitive Lab Report.pdf* o Provides detailed results of the braille cognitive lab	Most students take the online version of the CAI science assessment, but a paper/pencil version is available for students who need this accommodation. In addition, versions in Spanish and in Braille are also available. For the paper/pencil version, fixed forms were developed to meet each state's test blueprint. However due to the lack of enough appropriate items for paper/pencil delivery, some blueprint requirements of the DCI and PE levels were not met. (*CS 002-V2*, p26-27). • **States need to address this deficiency in their technical reports.** • CAI needs to provide details regarding plans for future item development aimed at the creation of more items for paper/pencil formatting to fill in the blueprint gaps that exist. It is not clear how the Spanish translations were conducted and whether the items were reviewed by native Spanish-speakers (*CS 002-V2*, p. 13) .
• (CE 4.6.2) Documented adequate evidence of comparability of the meaning and interpretations of the assessment results.	CE 4.6.2: • *CS 001 – V1_Shared Technical Report.pdf* o Section 1.4 – Provides item parameters for scoring (pp. 3-4; pdf pp. 8-9)	CE 4.6.2: It is unclear how the paper/pencil version of the assessment is scored, and therefore, evidence is not found to document that the scoring of the extended response items is consistent when comparing the online version and the paper/pencil versions of the assessment.

	● *CS 001.B – Science Calibration Item and Group Parameters.xlsx* ○ Reports item parameters used for operational scoring ● *CS 004.B – Braille Cognitive Lab Report.pdf* Provides detailed results of the braille cognitive lab	The many technical and administration issues raised by students in the Braille Cognitive Labs Study (*CS 004-V4*) would suggest that unless these issues have been addressed, the meaning and interpretation of test results for visually-impaired students would be suspect with regard to comparability with students who are not visually impaired. Evidence is not found regarding a comparability study between the Braille and the regular operational CAI assessment, nor is it clear if CAI has already rectified these issues or what their plan might be to do so. Evidence is needed to document the comparability of the Spanish-version, the ASL version, and the paper/pencil versions of the assessment, in comparison to the regular online version of the assessment **Participating states that administer the assessment using different types of devices (e.g., desktop computers, laptops, tablets) will need to provide evidence** to document that test-administration hardware and software (e.g., screen resolution, interface, input devices) are standardized across unaccommodated administrations; or (a) research reports (quantitative or qualitative) that show that variations resulting from different types of delivery devices do not alter the interpretations of results; or (b) an appropriate comparability study.

Section 4.6 Summary Statement

_ No additional evidence is required or

X__ The following additional evidence is needed/provide brief rationale:

- Evidence is needed to document the comparability of the Spanish-version, the ASL version, and the paper/pencil versions of the assessment, in comparison to the regular online version of the assessment.

Critical Element 4.7 – Technical Analysis and Ongoing Maintenance

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State: • (CE 4.7.1) Has a system for monitoring, 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), and	4.7.1: • *CS 001 – V1_Shared Technical Report.pdf* o Section 3.1.2 and Section 4.4 – Note when recommendations from states' technical advisory committees impacted the test development process (pp. 17-18; pdf pp. 22-23; p. 23; pdf p. 28) o Section 7 – Provides quality control procedures (pp. 41-42; pdf pp. 46-47) o Section 8 – Provides description of the Common Science Technical Advisory Committee • *CS 005 – V5_Shared Technical Report.pdf* o Section 3 – Provides quality monitoring system information (pp. 5-6; pdf pp. 7-8)	**Each participating state monitors the quality of its assessment system and will need to provide evidence of how they implement this state monitoring**. **State TACs provided input to help monitor aspects of the overall development process. States will need to provide evidence regarding the make-up of their state TACs.** CAI established a common Science TAC to make recommendations on technical and policy aspects of the CAI Science Assessment, including item and test construction, administration, scoring, and other issues. While the common TAC includes a representative from each of the participating state TACs, the overall make-up of the common TAC was not provided. It would be helpful for CAI to describe the balance of members by demographics and by expertise and/or role (e.g., assessment expert, psychometrician, NGSS/science expert, state/district administrator). CAI refers to a quality monitoring system (*CS 001-V1*, p 41) that yields item statistics, blueprint match rates, and item exposure rate reports. However, no evidence is found regarding the extent to which issues have been found, nor evidence regarding how this monitoring information is used for improving the quality of the assessment system over time. It is expected that states would be using these tools and providing follow-up. As a consortium, there should be some monitoring for the overall system for over exposure of items.
• (CE 4.7.2) Evidence of adequate technical quality is made public, including on the State's website.	4.7.2: See State-Specific Submissions.	

Section 4.7 Summary Statement
_X No additional evidence is required or ___ The following additional evidence is needed/provide brief rationale:

SECTION 5: INCLUSION OF ALL STUDENTS

Critical Element 5.1 – Procedures for Including Students with Disabilities

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or 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. Decisions about how to assess students with disabilities must be made by a student's IEP Team under IDEA, the placement team under Section 504, or the individual or team designated by a district to make that decision under Title II of the ADA, as applicable, based on each student's individual abilities and needs. If a State adopts alternate academic achievement standards for students with the most significant cognitive disabilities and administers an alternate assessment aligned with those standards under ESEA section 1111(b)(1)(E) and (b)(2)(D), respectively, the State must: • Establish guidelines for determining whether to assess a student with an AA-AAAS, including: o A State definition of "students with the most significant cognitive disabilities" that addresses factors related to cognitive functioning and adaptive behavior; • Provide information for IEP Teams to inform decisions about student assessments that:	See State-Specific Submissions.	

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
○ Provides a clear explanation of the differences between assessments aligned with grade-level academic achievement standards and those aligned with alternate academic achievement standards, including any effects of State and local policies on a student's education resulting from taking an AA-AAAS, such as how participation in such assessments may delay or otherwise affect the student from completing the requirements for a regular high school diploma; • Ensure that parents of students assessed with an AA-AAAS are informed that their child's achievement will be measured based on alternate academic achievement standards; • Not preclude a student with the most significant cognitive disabilities who takes an AA-AAAS from attempting to complete the requirements for a regular high school diploma; and • Promote, consistent with requirements under the IDEA, the involvement and progress of students with the most significant cognitive disabilities in the general education curriculum that is based on the State's academic content standards		

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
for the grade in which the student is enrolled; and • Develop, disseminate information on, and promote the use of appropriate accommodations to ensure that a student with the most significant cognitive disabilities who does not take an AA-AAAS participates in academic instruction and assessments for the grade in which the student is enrolled. • The State has in place and monitors implementation of guidelines for IEP teams to apply in determining, on a case-by-case basis, which students with the most significant cognitive disabilities will be assessed based on alternate academic achievement standards, if applicable. Such guidelines must be developed in accordance with 34 CFR § 200.6(d).[2]		

Section 5.1 Summary Statement

___ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:
- [list additional evidence needed w/brief rationale]

[2] See the full regulation at 34 CFR § 200.6(d) (online at https://www.ecfr.gov/cgi-bin/text-idx?SID=07e168e9e7a6c5931b4549cc15547ee9&mc=true&node=se34.1.200_16&rgn=div8)
Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 5.2 – Procedures for Including English Learners in Academic Content Assessments

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State has in place procedures to ensure the inclusion of all ELs in public elementary and secondary schools in the State's academic content assessments and clearly communicates this information to districts, schools, teachers, and parents, including, at a minimum: • Procedures for determining whether an EL should be assessed with a linguistic accommodation(s); • Information on accessibility tools and features available to all students and assessment accommodations available for ELs; • Assistance regarding selection of appropriate linguistic accommodations for ELs, including to the extent practicable, assessments in the language most likely to yield accurate and reliable information on what those students know and can do to determine the students' mastery of skills in academic content areas until the students have achieved English language proficiency.	See State-Specific Submissions.	

Section 5.2 Summary Statement

___ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:
- [list additional evidence needed w/brief rationale]

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 5.3 – Accommodations

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State makes available appropriate accommodations and ensures that its assessments are accessible to students with disabilities and ELs, including ELs with disabilities. Specifically, the State: • (CE 5.3.1) Ensures that appropriate accommodations, such as, interoperability with, and ability to use, assistive technology, are available to measure the **academic achievement** of students with disabilities.	CE 5.3.1: • *CS 001 – V1_Shared Technical Report.pdf* o Section 1.4 – Provides available test formats and special versions (pp. 3-4; pdf pp. 8-9) o Section 2.3 – Outlines the designated supports and accommodations available (pp. 5-6; pdf pp. 10-11) • *CS 002 − V2_Shared Technical Report.pdf* o Section 1.2 – Discusses the underlying principles guiding development and specifies that CAI's test delivery system has received WCAG 2.0 AA certification, offers a wide range of accessibility tools, and is compatible with most assistive technologies (pp. 2-3; pdf pp. 5-6) o Section 2.5.4 – Discusses text-to-speech tagging, Spanish translations and braille (p. 13; pdf p. 16) o Section 4.4 – Provides paper-pencil accommodations form construction (pp. 26-27; pdf pp. 30-31) • *CS 005 – V5_Shared Technical Report.pdf* o Section 2.1.4 – Lists brochures and quick guides (p. 4; pdf p. 6) o Section 5 – Discusses online testing features and accommodations (pp. 7-13; pdf pp. 9-15)	CE 5.3.1: Evidence is provided that appropriate accommodations, including assistive technology, are available to measure the science achievement of students with disabilities. **States will need to provide evidence that:** • **they are using the same accommodations for students with disabilities as available through CAI, and if not, an explanation regarding their options for accommodations for students with disabilities, as well as a summary of the frequency of use of each accommodation.** • **there is a process to individually review exceptional requests and ensure that accommodations do not deny students with disabilities the opportunity to participate.**
• (CE 5.3.2) Ensures that appropriate accommodations are available for ELs;	CE 5.3.2: • *CS 002 − V2_Shared Technical Report.pdf* o Section 1.2 – Discusses the underlying principles guiding development (p. 2; pdf p. 5)	CE 5.3.2: Evidence is provided that appropriate accommodations are available to measure the science achievement of English Learners. **States will need to provide evidence that:**

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
	○ Section 2.5.4 – Discusses text-to-speech tagging, Spanish translations and braille (p. 13; pdf p. 16) ○ Section 4.4 – Provides paper-pencil accommodations form construction (pp. 26-27; pp. 30-31) • *CS 005 – V5_Shared Technical Report.pdf* ○ Section 2.1.4 – Lists brochures and quick guides (p. 4; pdf p. 6) ○ Section 5 – Discusses online testing features and accommodations (pp. 7-13; pdf pp. 9-15)	• **they are using the same accommodations for EL students as available through CAI, and if not, an explanation regarding their options for accommodations for EL students, as well as a summary of the frequency of use of each accommodation.** • **there is a process to individually review exceptional requests and ensure that accommodations do not deny EL students the opportunity to participate.**
• (CE 5.3.3) Has determined that the accommodations it provides (1) are appropriate and effective for meeting the individual student's need(s) to participate in the assessments, (2) do not alter the construct being assessed, and (3) allow meaningful interpretations of results and comparison of scores for students who need and receive accommodations and students who do not need and do not receive accommodations;	CE 5.3.3: • *CS 002 − V2_Shared Technical Report.pdf* ○ Section 1.2 – Discusses the underlying principles guiding development (p. 2; pdf p. 5) ○ Section 2.5.4 – Discusses text-to-speech tagging, Spanish translations and braille (p. 13; pdf p. 16) ○ Section 4.4 – Provides paper-pencil accommodations form construction (pp. 26-27; pp. 30-31) • *CS 001 – V1_Shared Technical Report.pdf* ○ Section 3.1.2 – Discusses Differential Item Functioning for the 2019 Field Test (pp. 14-18; pdf pp. 19-23) ○ Section 4.4– Analyzes Differential Item Functioning (pp. 21-23; pdf pp. 26-28) • *CS 005 – V5_Shared Technical Report.pdf* ○ Section 5 – Discusses how states determine appropriateness and provide online testing features and accommodations (p. 7-13; pdf p. 9-15)	CE 5.3.3: While CAI states that "State-approved accommodations do not compromise the learning expectations, constructs, or grade-level standards" and "increase the validity of inferences about students with disabilities" (*CS V1*, p 4), no evidence is found for substantiating these claims. It is not clear if CAI conducted studies to ensure the validity of these claims, or if they are relying on the fact that the accommodations included in the CAI assessment are the common accommodations included in many large-scale state assessments. As noted in other sections of these notes, the following observations bring into question whether meaningful interpretations of results for the following groups can be assured: ○ For students who take the paper/pencil form of the test, no evidence is provided regarding a comparability study of the scoring of extended response items between these students and those who take the online version of the test. ○ For Spanish-speaking ELs, no evidence is provided regarding the review of the Spanish translations by native Spanish speakers.

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
		○ For visually-impaired students, the Braille version suffered from a number of formatting and technical issues that bring into question whether the results for these students are comparable to results for students not taking the Braille version. **States will need to provide evidence whether work was conducted in their state regarding the above concerns.**
• (CE 5.3.4) Has a process to individually review and allow exceptional requests for a small number of students who require accommodations beyond those routinely allowed. • (CE 5.3.5) Ensures that accommodations for all required assessments do not deny students with disabilities or ELs the opportunity to participate in the assessment and any benefits from participation in the assessment.	CE 5.3.4: See State-Specific Submissions. CE 5.3.5: See State-Specific Submissions.	CE 5.3.4: See State-Specific Submissions. CE 5.3.5: See State-Specific Submissions.

Section 5.3 Summary Statement

___ No additional evidence is required or

_X The following additional evidence is needed/provide brief rationale:

- Evidence to support the claim that the accommodations used on the assessment "do not compromise the learning expectations, constructs, or grade-level standards" and "increase the validity of inferences about students with disabilities" as well as EL students.

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 5.4 – Monitoring Test Administration for Special Populations

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State monitors test administration in its districts and schools to ensure that appropriate assessments, with or without accommodations, are selected for all students with disabilities and ELs 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 under IDEA, placement team convened under Section 504; or for students covered by Title II of the ADA, the individual or team designated by a district to make these decisions; or another process for an EL; • Administered with fidelity to test administration procedures; • Monitored for administrations of all required academic content assessments and AA-AAAS.	See State-Specific Submissions	

Section 5.4 Summary Statement

___ No additional evidence is required or

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
___ The following additional evidence is needed/provide brief rationale: • [list additional evidence needed w/brief rationale]		

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

SECTION 6: ACADEMIC ACHIEVEMENT STANDARDS AND REPORTING

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

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
For academic content standards: The State formally adopted **challenging academic achievement standards** in reading/language arts, mathematics, and science for all students, specifically: • The State formally adopted academic achievement standards in the required tested grades and, at its option, alternate academic achievement standards for students with the most significant cognitive disabilities; • The State applies its 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: (1) at least three levels of achievement, with two for high achievement and a third for lower achievement; (2) descriptions of the competencies associated with each achievement level; and (3) achievement scores that differentiate among the achievement levels.	See State-Specific Submissions	

Section 6.1 Summary Statement

___ No additional evidence is required or

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
___ The following additional evidence is needed/provide brief rationale: • [list additional evidence needed w/brief rationale]		

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 6.2 – Achievement Standards Setting

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State used a technically sound method and process that involved panelists with appropriate experience and expertise for setting: • **Academic achievement standards and, as applicable, alternate academic achievement standards**.	• *CS 003 – V3_Shared Technical Report.pdf* o Section 2 – Provides the purpose and overview for the common standard-setting methodology (p.1; pdf p. 4) o Section 5 – Provides the standard-setting process (pp. 3-19; pdf pp. 7-26) o Section 5.4 – Describes the panelists and their role (p. 8; pdf p. 11) o Section 5.11 – Provides the evaluation process used to measure procedural validity (pp. 19-20; pdf pp. 22-23) o Section 6.1 – Describes the adherence to AREA/APA/NCME standards (p. 20; pdf p. 23) • *CS003.A – Development of Range Performance-Level Descriptors.pdf* o Describes the system of performance-level descriptors and the process for developing range performance-level descriptors • *CS003.B – Standard-Setting Training Slides.pdf* o Describes all steps, roles and responsibilities in standard-setting • *CS003.C – Standard-Setting Practice Quiz.pdf* o Includes quiz for determining if standard-setting panelists understand the process of standard-setting • *CS003.D – Standard-Setting Readiness Forms.pdf* o Includes forms that allow workshop facilitators to determine if panelists are ready to proceed with the standard-setting process • *CS004.E − Appendix E. Synopsis of Validity Evidence for the Cut Scores.pdf* o External review of the standard-setting method and implementation by independent measurement expert	CAI established a technically-sound process for states to use in establishing their own cut scores, performance levels, and content-based performance level descriptors that reflect the state's content standards. Cambium developed a new methodology for standard setting, but it follows similar protocols for other methods and does not vary substantially from Skaggs' passage-based bookmark method. **States will need to provide evidence that they followed the process provided by CAI, a description of the panelists and resources used in setting their performance standards, any deviations they may have embedded in their processes and the rationale for such decisions.** It would be helpful to see documentation regarding recommendations from Cambium to states for the appropriate number and distribution of panelists for state standard-setting, and to especially note the importance of including panelists with understanding of and experience with using three-dimensional science standards. **States will need to provide evidence on the expertise of panelists in their state standard-setting, as well as documentation that the training given to their panelists includes a focus on three-dimensional science standards.** It would be helpful for CAI to provide their rationale for using the range PLDs based on the standards drafted by the Washington State Office of Superintendent of Public Instruction (OSPI) as a starting point, as noted in *CS 003 - V3_Shared Technical Report*.

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
		The alignment among blueprint specs, item difficulties, and PLDs may be suspect (e.g., there is no indication of conservation of mass in level 2 on Table 2, but it is included in levels 3 and 4) as described in *CS 003.A - Appendix A. Development of Range Performance-Level Descriptors*. **This may be an issues for states to address in CE 6.3.** **States will need to address the validity of the process in terms of panelists' evaluation of the standard-setting workshop.**

___ No additional evidence is required or

_X_ The following additional evidence is needed/provide brief rationale:

- Evidence that each panel includes at least one NGSS expert and that training provides a strong focus on three-dimensional standards, to help ensure that alignment to the NGSS is met at all levels for all three dimensions and not simply in the area of science content.

Critical Element 6.3 – Challenging and Aligned Academic Achievement Standards

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
For academic achievement standards: The State's academic achievement standards are challenging and aligned with the State's academic content standards and with entrance requirements for credit-bearing coursework in the system of public higher education in the State and relevant State career and technical education standards such that a 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 they graduate from high school in order to succeed in college and the workforce. If the State has adopted alternate academic achievement standards for students with the most significant cognitive disabilities, the alternate academic achievement standards (1) are aligned with the State's challenging academic content standards for the grade in which a student is enrolled; (2) promote access to the general curriculum consistent with the IDEA; (3) reflect professional judgment as to the highest possible standards achievable for such students; (4) are designated in the IEP for each student for whom alternate academic achievement standards apply; and (5) are aligned to ensure that a student who meets the alternate academic achievement standards is on track to pursue	See State-Specific Submissions	

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
postsecondary education or competitive integrated employment.		
Section 6.3 Summary Statement		
___ No additional evidence is required or ___ The following additional evidence is needed/provide brief rationale: • [list additional evidence needed w/brief rationale]		

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element 6.4 – Reporting

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State reports its assessment results for all students assessed, and the reporting facilitates timely, appropriate, credible, and defensible interpretations and uses of those results by parents, educators, State officials, policymakers and other stakeholders, and the public. The State reports to the public its assessment results on **student academic achievement for all students and each student group at each achievement level**[3] For **academic content assessments**, 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 academic content assessments that:	See State-Specific Submissions	

[3] Although all students with disabilities must be included in a State's assessment system, requirements for public reporting in ESEA section 1111(h)(1)(C)(ii) apply only to children with disabilities as defined in section 602(3) of the IDEA.

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
○ Provide valid and reliable information regarding a **student's academic achievement**; ○ Report the **student's academic achievement** in terms of the State's grade-level academic achievement standards; ○ Provide information to help parents, teachers, and principals interpret the test results and address the specific **academic needs of students**; ○ Are provided in an understandable and uniform format; ○ Are, to the extent practicable, written in a language that parents and guardians can understand or, if it is not practicable to provide written translations to a parent or guardian with limited English proficiency, are orally translated for such parent or guardian; ○ Upon request by a parent who is an individual with a disability as defined by the ADA, as amended, are provided in an alternative format accessible to that parent. • 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.		

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
Section 6.4 Summary Statement		
___ No additional evidence is required or ___ The following additional evidence is needed/provide brief rationale: • [list additional evidence needed w/brief rationale]		

SECTION 7: LOCALLY SELECTED NATIONALLY RECOGNIZED HIGH SCHOOL ACADEMIC ASSESSMENTS
(**if applicable**; evidence for this section would be submitted in ADDITION to evidence for sections 1 through 6)

Critical Element 7.1 – State Procedures for the Use of Locally Selected, Nationally Recognized High School Academic Assessments

Critical Element	**Evidence (Record document and page # for future reference)**	**Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence**
The State has established technical criteria to use in its review of any submission of a locally selected, nationally recognized high school academic assessment. The State has completed this review using its established technical criteria and has found the assessment meets its criteria **prior to** submitting for the Department's assessment peer review. The State's technical criteria include a determination that the assessment: • Is aligned with the challenging State academic standards; and • Addresses the depth and breadth of those standards. AND The State has procedures in place to ensure that a district that chooses to use a nationally recognized high school academic assessment administers the same assessment to all high school students in the district except for students with the most significant cognitive disabilities who may be assessed with an AA-AAAS.	N/A	

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
AND The technical criteria established by the State in reviewing a locally selected, nationally recognized high school academic assessment must ensure that the use of appropriate accommodations does not deny a student with a disability or an EL— • The opportunity to participate in the assessment; and • Any of the benefits from participation in the assessment that are afforded to students without disabilities or students who are not ELs.		

Section 7.1 Summary Statement

___ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:
- [list additional evidence needed w/brief rationale]

Element 7.2 –State Monitoring of Districts Regarding the Use of Locally Selected, Nationally Recognized High School Academic Assessments

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The State must have procedures in place to ensure that: **Before** a district requests approval from the State to use a nationally recognized high school academic assessment, the district notifies all parents of high school students it serves— • That the district intends to request approval from the State to use a nationally recognized high school academic assessment in place of the statewide academic assessment; • Of how parents and, as appropriate, students may provide meaningful input regarding the district's request (includes students in public charter schools who would be included in such assessments); and • Of any effect of such request on the instructional program in the district.	N/A	

Section 7.2 Summary Statement

___ No additional evidence is required or

___ The following additional evidence is needed/provide brief rationale:
- [list additional evidence needed w/brief rationale]

Consistent with the note on page 1, the evidence requested by the peer reviewers does not necessarily reflect 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 for the assessment peer review. As a result, a State should refer to the letter to the State, including the list of additional evidence needed, if any, from the Department.

Element 7.3 –Comparability of the Locally Selected Nationally Recognized High School Academic Assessments with the State Assessments

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
The locally selected, nationally recognized high school academic assessment: • Is equivalent to or more rigorous than the statewide assessment, with respect to— o The coverage of academic content; o The difficulty of the assessment; o The overall quality of the assessment; and o Any other aspects of the assessment that the State may establish in its technical criteria; • Produces valid and reliable data on student academic achievement with respect to all high school students and each subgroup of high school students in the district that— o Are comparable to student academic achievement data for all high school students and each subgroup of high school students produced by the statewide assessment at each academic achievement level; o Are expressed in terms consistent with the State's academic achievement standards; and o Provide unbiased, rational, and consistent differentiation among schools within the State for the purpose of the State determined accountability system including calculating the Academic Achievement indicator and annually meaningfully differentiating between schools.		

Critical Element	Evidence (Record document and page # for future reference)	Comments/Notes/Questions/Suggestions Regarding State Documentation or Evidence
Section 7.3 Summary Statement		
___ No additional evidence is required or ___ The following additional evidence is needed/provide brief rationale: • [list additional evidence needed w/brief rationale]		