Proviso Township High School District 209

Curriculum Management Audit

Action Plan for Priority Improvement Initiatives

Review dates October 2022 - April 2023

Prepared by Tim Suba

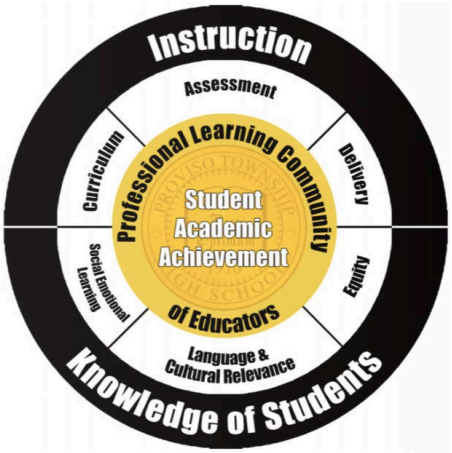
Insight EDUCATION GROUP

Table of Contents

Internal System Coherence

ALL ABOARD

Equity, Excellence, Relevance for ALL.
Pursuit of Nothing But The Best



The Curriculum Management Audit aims to assess systems of curriculum and assessment within Proviso Township High School District 209 (PTHS D209). Essentially, the objective of the Audit is to gain a deeper understanding of the current state of curriculum systems and materials, as well as perspectives of various stakeholder groups relative to those systems and materials, in order to suggest next steps that would support key aspects of the district's Instructional Framework. These next steps, in turn, seek to advance the five goals of the district's Strategic Plan.

The Audit proceeded in three stages--Curriculum Review, Gap Analysis, and Action Plan for Priority Improvement Initiatives.

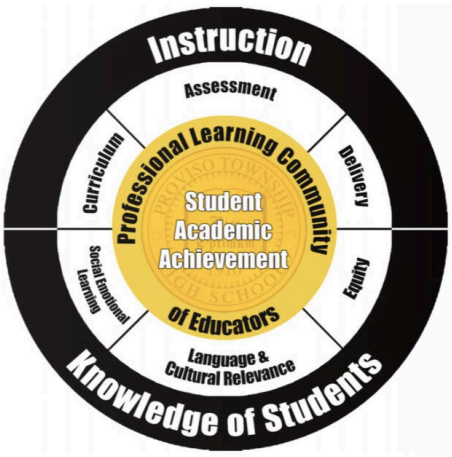
The goal of the Curriculum Review stage was to understand what curriculum materials were available and being used for instruction across the district. The Insight team collected both district-level and school level quantitative and qualitative data. The data consisted of a survey and reviews of curriculum artifacts using indicators from leading curriculum analysis tools. A synthesis of the data can be found in the pages that follow.

During the Gap Analysis stage, the Insight team sought to understand what barriers existed, if any, between present conditions and ideal outcomes. During this stage, Insight conducted focus group interviews with several different stakeholder groups, including groups of teachers, coordinators, school leaders, and students.

Finally, the data and evidence gathered are used to generate this report, the Action Plan for Priority Improvement Initiatives. This report is organized according to the data collection opportunities, followed by identification of bright spots (or practices already present), areas of opportunity, then suggested next steps for the district.

We hope the findings and recommendations in this report can serve as a springboard to the planning phase. We are so grateful for the opportunity to partner with you and look forward to supporting you on this journey in service of students.

ALIGNMENT TO STRATEGIC PLAN AND INSTRUCTIONAL FRAMEWORK



The district's Instructional Framework states that "in order to increase student academic achievement," the district "is committed to high-quality instruction through a guaranteed, viable **curriculum**, ongoing **assessment**, and high-impact instructional **delivery**."

The Curriculum Management Audit seeks to examine curriculum systems and materials (including assessments). Curriculum and assessments, in turn, impact delivery. High-quality curriculum, assessment, and delivery, all driven by Proviso's dedicated team of educators, are factors that contribute to the success of students.

Support of the Instructional Framework aims to advance all five goals of the district's Strategic Plan. The goals are:

The next pages outline the alignment between the Audit and each of these five goals.

ALIGNMENT TO STRATEGIC PLAN AND INSTRUCTIONAL FRAMEWORK

The Audit supports this goal's objective of "support[ing] a culture of continuous improvement" by contributing data that will sustain conversations across PTHS D209 on matters concerning district-wide improvement.

As mentioned previously, the Audit's process and findings relate to the statements in the Instructional Framework. That said, this goal's objective is that "all students will have access to and engagement with high-quality, equitable, and relevant instruction." The Audit's recommendations will help to positively impact curriculum and assessment, and therefore, delivery.

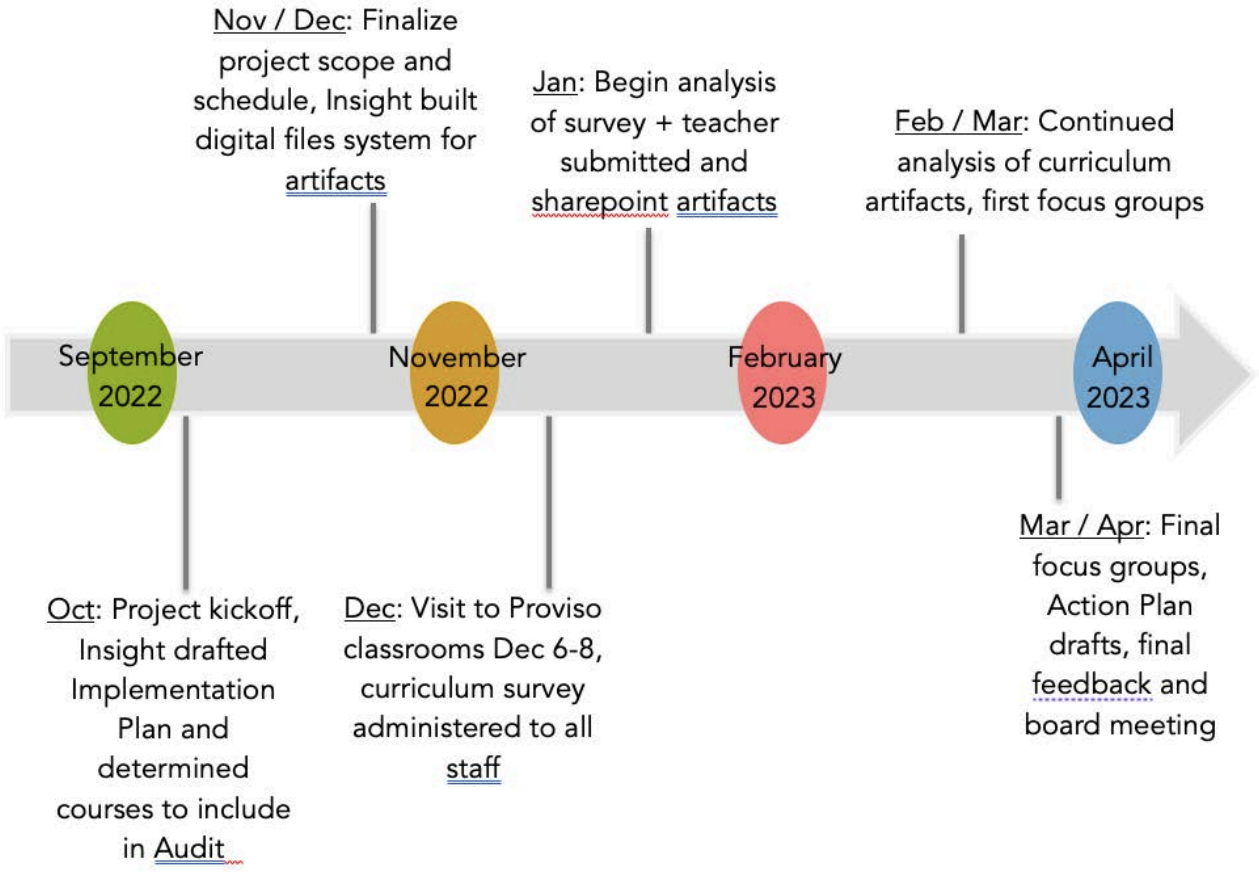
ALIGNMENT TO STRATEGIC PLAN AND INSTRUCTIONAL FRAMEWORK

The Audit's next steps around curriculum materials and professional development support the goal's objective to "design, implement, and sustain a technology plan" that will allow "integration of technology into teaching and learning."

As one of this goal's Strategic Action Steps is to "develop a system-wide process for improving the capacity of teachers and administrators" through professional development, this Audit suggests systems of ongoing collaboration among staff, and ongoing professional development to drive high-quality, equitable, and relevant instruction.

"Efficient operations that support the instructional core" drive a number of this Audit's recommendations.

TIMELINE OF EVENTS



Curriculum Review

The Audit's Curriculum Review stage focused on a comprehensive collection of courses based proportionally on PTHS D209's two sets of graduation requirements. Data from documents for review, classroom visits, focus groups (comprising teachers, school leaders, or support staff), and surveys were selected from the courses below:

English (6 courses)
English I
English I EL
English II
AP English Language and Composition
IH Survey of Literature
IH World Literature

Mathematics (5 courses)
Integrated Math I
Integrated Math I Instructional
Integrated Math III
AP Calculus AB
IH Integrated Math I

Science (4 courses)
Biology
IH Biology
Chemistry
IH Chemistry

Social Studies (3 courses)
US History
IH US History
IH Global Studies

Health/Wellness (3 courses)
Sophomore PE
Driver's Education
IH Driver's Education

World Languages (2 courses)
Spanish I
IH Spanish I

These courses are a beginning representation of PTHS D209's diverse offerings and diverse student body; limiting data to that obtained from these courses allows the Audit to go deeper with fewer representative courses, and thus serves the intended goals of this Audit.

Insight conducted focus groups and administered surveys to, teachers, leaders, and other staff who teach or are otherwise impacted by the courses above. The focus groups and surveys will attempt to ascertain baseline curriculum and content practices across these courses.

Finally, Insight gathered documents (curriculum documents, assessment and instructional materials, syllabi, calendars, etc.) related to the courses above.

CLASSROOM VISITS

From December 6 to December 8, 2022, the Insight team conducted classroom visits at each of PTHS D209's three high schools. The classes visited were as follows:

	English	Mathematics	Science
Proviso East HS	English I (2 classrooms), English I EL	Integrated Math I, Integrated Math III, AP Calculus AB	Biology (2 classrooms), Chemistry (2 classrooms)
Proviso West HS	English I EL, English II	Integrated Math I (2 classrooms), Integrated Math III	Biology (2 classrooms), Chemistry (2 classrooms)
PMSA	IH Survey of Literature (2 classrooms), IH World Literature	IH Integrated Math I, AP Calculus AB	IH Biology, IH Chemistry
	Social Studies	**Health/Wellness**	**World Languages**
Proviso East HS	US History (2 classrooms, IH Global Studies	Driver's Education, Sophomore PE	Spanish I
Proviso West HS	US History	Driver's Education, Sophomore PE	Spanish I
PMSA	IH US History	IH Driver's Education, IH Wellness II	IH Spanish I

Insight also visited a Consumer Education class at Proviso East, an Art Foundations class at Proviso West, and IH Integrated Math III and IH Visual Arts classes at PMSA. However, those courses were ultimately not included in the Audit.

CLASSROOM VISITS

Insight's objective for the visits was to see a school day in action. During each ten-minute visit, the Insight team experienced classroom instruction, noting lesson objective(s), the Illinois Learning Standard(s) (or other course standards) addressed, curriculum materials used, and intended student product. The information gathered provided background and context to the curriculum artifacts (described in the next section of this report) later collected and examined.

Insight also spoke with teachers and answered any questions (time and opportunity permitting), and connected with school leaders.

Insight thanks PTHS D209's Educational Services team, school leaders, and of course, the teachers of Proviso East, Proviso West, and PMSA for their coordination before, and hospitality during, the visits.

SURVEY DATA

A vital component of the process included collecting and analyzing survey data from teachers across all three campuses. The survey was administered in December of 2022 and asked respondents to provide information on key curriculum issues. For example, do they have access to district-provided curriculum and resources? If so, how frequently do they use those resources (if they exist)?

KEY FINDINGS

- There were 208 responses. 29% of respondents reported not having, or not knowing if they had, access to district-provided curriculum materials.
- Slightly less than half (46%) reported not having the teacher resources they needed to successfully teach their course(s).
- A common issue reported across all content areas was inconsistent access to (and usage of) common resources, like unit plans and curriculum maps, on district share drives.

ARTIFACT ANALYSIS

As mentioned previously, the Curriculum Management Audit focused on a limited number of courses throughout the district. Teachers were randomly selected to contribute to the overall curriculum picture at PTHS D209 by providing artifacts for one of the courses in the Audit.

The courses, per school, were:

	Math	English	Science	Social Studies	Health / Wellness	World Languages
Proviso East	AP Calculus AB, Integrated Math I (2 classrooms), Integrated Math III	AP English Language and Composition, English I (2 classrooms), English II	Biology (2 classrooms), Chemistry (2 classrooms)	US History (2 classrooms), IH Global Studies	Driver's Education, Sophomore PE	Spanish I
Proviso West	Integrated Math I (2 classrooms), Integrated Math III	AP English Language and Composition, English I (2 classrooms), English II	Biology (2 classrooms), Chemistry (2 classrooms)	US History (2 classrooms), IH Global Studies	Driver's Education, Sophomore PE	Spanish I
PMSA	AP Calculus AB, IH Integrated Math I (2 classrooms)	AP English Language and Composition, IH Survey of Literature (2 classrooms), IH World Literature	IH Chemistry, IH Biology	IH US History, IH Global Studies	IH Driver's Ed, IH Health & Wellness II	IH Spanish I

Artifacts included end-of course assessments, unit assessments, and shorter formative assessments. Artifacts were assessed on criteria derived from assessment evaluation tools used by field-leading organizations who have influenced the creation and direction of the Common Core State Standards (and thus the Illinois Learning Standards).

One hundred eleven artifacts were collected across the three schools—either through teacher submission or selection from Sharepoint. The number and variety of artifacts, by content area, is shown below:

	Math	English	Science	Social Studies	Health / Wellness	World Languages
End-of-Course Assessments	2	0	8	4	2	1
End-of-Unit or Common Interim Assessments	12	3	16	6	3	5
Shorter Formative Assessments	11	11	12	4	6	5
Total	25	14	36	14	11	11

Insight thanks PTHS D209's Educational Services for their coordination in this effort, and course, the teachers of Proviso East, Proviso West, and PMSA for their efforts in submitting materials.

The following pages show artifact analysis by content, along with summarized findings from the artifacts. Detailed findings are in the appendix.

MATHEMATICS

Insight used criteria derived from leading assessment evaluation tools and appropriate for the purposes of this Audit, to assess math assessments on a 3-point scale. Evaluation tools of origin are linked.

Five Math Assessment Criteria and 3-point scale

A. The set of items is clearly consistent with the most important content of the identified standard, and items should be designed to elicit direct, observable evidence of a student's ability to independently demonstrate competency (*from EQuIP Rubric, achieve.org*)
B. Item set is consistent with the standards' primary aspect of rigor (conceptual, procedural, and/or application) (*from Criteria for Procuring and Evaluating High-Quality Assessments, CCSSO*)
C. Assessments should contain a variety of item types (*from Criteria for Procuring and Evaluating High-Quality Assessments, CCSSO*)
D. Assessment should demonstrate authentic connections between the content standards and the eight Standards for Mathematical Practices (*from Assessment Evaluation Tool, achievethecore.org*)
E. Majority of items on the assessment come from major work of the grade (priority standards) (*from Assessment Evaluation Tool, achievethecore.org*)

3	a significant majority or all aspects of the criterion
2	some aspects of the criterion
1	none or nearly none of the aspects of the criterion

For shorter artifacts, Insight used the EQuIP Task Review Rubric for Mathematics, also used by ISBE. It assesses alignment to standards, attention to the instructional shifts in the standards, and implementation support. Insight rated these artifacts with the following descriptors:

E	Most criteria checked
E/I	Many criteria checked but could use minor improvements
R	Some criteria checked
N	Task not recommended for instruction

MATHEMATICS

Seventeen math assessments were rated on the five Math Assessment Criteria:

Criteria	Average score (between 1-3)
A. The set of items is clearly consistent with the most important content of the identified standard, and items should be designed to elicit direct, observable evidence of a student's ability to independently demonstrate competency.	2.29
B. Item set is consistent with the standards' primary aspect of rigor (conceptual, procedural, and/or application).	2.35
C. Assessments should contain a variety of item types.	1.88
D. Assessment should demonstrate authentic connections between the content standards and the eight Standards for Mathematical Practices.	1.82
E. Majority of items on the assessment come from major work of the grade (priority standards).	1.82

Eight shorter artifacts were rated on the EQuIP Task Review Rubric:

Descriptor	Number of Artifacts
E: Most criteria checked	4
E/I: Many criteria checked but could use minor improvements	1
R: Some criteria checked	3
N: Task not recommended for instruction	0

MATHEMATICS

FINDINGS

1. Assessments with strongest variety and standard alignment were those submitted directly from high-quality curriculum materials sources, especially including Savvas enVision Integrated Mathematics (2019).

2. AP Calculus AB common interim assessments had item variety consistent with the 50/50 selected response-to-open response split. Later common interims (e.g., CID or CIE) should be closer to the length of the AP exam, with more than one open response item.

3. Some summative (or common interim) assessments are built around only one standard, which accounts for some of the lower scores on criteria A, B, and E.

4. Some teacher-created assessments and Quizizz tasks containing mostly selected response (multiple-choice) needed more item variety and were mostly procedural items, thus scoring lower on criteria B, C, and D.

5. Strongest assessments were aligned to standards, required students to use Standards for Mathematical Practice (e.g., looking for structure and making use of it), and had overall balance and integration of rigor (conceptual, application, and procedural). Items also connected to observable parts of the standard.

6. Strongest shorter formative assessments were comprehensive, focused, and also aligned to standards. These formative assessments briefly tapped into prerequisite knowledge, asked students to explain why and check for reasonableness, and asked them to explain procedures in their own words, in addition to solving them.

7. Only two final exams were submitted; end-of-course exams (besides those for AP course) do not appear to be standardized across classes, across schools.

ENGLISH

Insight used criteria derived from leading assessment evaluation tools and appropriate for the purposes of this Audit, to assess English assessments on a 3-point scale. Evaluation tools of origin are linked.

Five English Assessment Criteria and 3-point scale

A. Texts are worth reading. Should include high quality texts at appropriate lexile and qualitative complexity (*from Criteria for Procuring and Evaluating High-Quality Assessments, CCSSO*)
B. Balance of genres when appropriate; reflect demands of shifts (citing evidence, building knowledge, text complexity) and standards (*from Criteria for Procuring and Evaluating High-Quality Assessments, CCSSO*)
C. Assessments and tasks should include purposefully linked passages or questions that allude to other accessible texts or passages according to RI/RL.9 (*from Assessment Evaluation Tool, achievethecore.org*)
D. Variety of item types to accurately assess a standard. Students should have the opportunity to write in response to high-quality texts (*from Criteria for Procuring and Evaluating High-Quality Assessments, CCSSO*)
E. Items should be designed to elicit direct, observable evidence of the degree to which a student can independently demonstrate the key understandings of the text relative to the demands of the standard (*from EQuIP rubric, achieve.org*)

3	a significant majority or all aspects of the criterion
2	some aspects of the criterion
1	none or nearly none of the aspects of the criterion

For shorter artifacts, Insight used the EQuIP Task Review Rubric for ELA, also used by ISBE. It assesses standard alignment, attention to text complexity and the shifts in the standards, and implementation support. Insight rated these artifacts with the following descriptors:

E	Most criteria checked
E/I	Many criteria checked but could use minor improvements
R	Some criteria checked
N	Task not recommended for instruction

ENGLISH

Eleven English assessments were rated on the five English Assessment Criteria:

Criteria	Average score (between 1-3)
A. Texts are worth reading. Should include high quality texts at appropriate lexile and qualitative complexity.	2.45
B. Balance of genres when appropriate; reflect demands of shifts (citing evidence, building knowledge, text complexity) and standards.	1.55
C. Assessments and tasks should include purposefully linked passages or questions that allude to other accessible texts or passages according to RI/RL.9.	1
D. Variety of item types to accurately assess a standard. Students should have the opportunity to write in response to high-quality texts.	2.09
E. Items should be designed to elicit direct, observable evidence of the degree to which a student can independently demonstrate the key understandings of the text relative to the demands of the standard.	2.45

Three shorter artifacts were rated on the EQuIP Task Review Rubric:

Descriptor	Number of Artifacts
E: Most criteria checked	0
E/I: Many criteria checked but could use minor improvements	1
R: Some criteria checked	2
N: Task not recommended for instruction	0

ENGLISH

FINDINGS

1. English I and English II common interim assessments all featured texts worth reading (publishable quality and/or from published sources, quantitative complexity with grade level-appropriate lexile levels, and qualitative complexity with grade level-appropriate themes and subjects).

2. AP English Language and Composition common interim assessments had mostly selected response, but AP exam only has 45% selected response. Item types in common interims should be closer ratio found on the AP exam.

3. Many assessments did not have a text to read and to which to respond, or featured texts which were not grade level-appropriate, which accounts for some of the lower scores on criteria A, C, and E.

4. Some formative assessments had questions worth answering or exploring, but lacked anchor or suggested grade-level texts to which students should refer or cite, thus scoring lower on criteria A, B, and E.

5. Strongest assessments were aligned to standards, required students to a fiction and nonfiction text, and variety of standards. Items also connected to observable parts of the standard.

6. Some formative assessments used grade level-appropriate texts, but had prompts that mostly asked students to identify and recall. Enrich these prompts by pushing students to explore themes and essential questions more deeply with writing, all while requiring them to cite evidence from text.

7. No final exams were submitted; end-of-course exams (besides those for AP courses) do not appear to be standardized across classes, across schools.

SCIENCE

Insight used criteria derived from various EQuIP/NGSS science rubrics from Achieve and appropriate for the purposes of this Audit, to assess science assessments on a 3-point scale.

Five Science Assessment Criteria and 3-point scale

A. Assesses state science standards to provide evidence about students' achievement in science. Assessment requires students to use some understanding of Disciplinary Core Ideas to successfully complete it, and includes Reading and Writing for Science and Technical standards.
B. Assessment requires students to use at least one Science and Engineering Practice to successfully complete the task.
C. Assessment requires students to identify and interpret evidence and engage in scientific reasoning as they make sense of phenomena and address problems.
D. There are varied task types requiring a range of analytical thinking and cognitive complexity.
E. Majority of assessment cannot be answered without information from tasks or items, nor can the majority of the assessment's items be answered successfully by using rote knowledge.

3	a significant majority or all aspects of the criterion
2	some aspects of the criterion
1	none or nearly none of the aspects of the criterion

For labs or shorter artifacts, Insight adapted the EQuIP Task Review Rubrics, also used by ISBE. This rubric assesses standard alignment, attention cross-cutting concepts and science and engineering practices, and implementation support. Insight rated these artifacts with the following descriptors:

E	Most criteria checked
E/I	Many criteria checked but could use minor improvements
R	Some criteria checked
N	Task not recommended for instruction

SCIENCE

Thirty science assessments were rated on the five science Assessment Criteria:

Criteria	Average score (between 1-3)
A. Assesses state science standards to provide evidence about students' achievement in science. Assessment requires students to use some understanding of Disciplinary Core Ideas to successfully complete it, and includes Reading and Writing for Science and Technical standards.	2.1
B. Assessment requires students to use at least one Science and Engineering Practice to successfully complete the task.	2.07
C. Assessment requires students to identify and interpret evidence and engage in scientific reasoning as they make sense of phenomena and address problems.	1.47
D. There are varied task types requiring a range of analytical thinking and cognitive complexity.	1.77
E. Majority of assessment cannot be answered without information from tasks or items, nor can the majority of the assessment's items be answered successfully by using rote knowledge.	1.77

Six labs or artifacts were rated on the adapted EQuIP Task Review Rubric:

Descriptor	Number of Artifacts
E: Most criteria checked	2
E/I: Many criteria checked but could use minor improvements	2
R: Some criteria checked	1
N: Task not recommended for instruction	1

SCIENCE

FINDINGS

1. Assessments scoring higher on the criteria featured data sets and figures that required analysis to succeed on the assessment (vs. using rote or memorized knowledge). Assessments featuring this type of analysis require students to use more cross-cutting concepts and science and engineering practices (as defined by the NGSS) to succeed.

2. Many assessments contained mostly selected response, which accounts for some of the lower scores on all criteria.

3. Strongest assessments were aligned to standards and had strong item variety (including Open Response). These assessments asked students to write, interpret graphs and charts, and use problem solving skills using data presented.

4. Strongest labs included opportunities for students to analyze data, argue from evidence, and draw conclusions. Prompts from these labs encouraged discovery, in which students might first be asked to write or draw what was observed, followed by questions prompting students to connect observations, through reasoning with evidence, to phenomena studied or discussed.

5. There is more potential to include more cross-cutting concepts and science and engineering practices in assessments and tasks by including fewer rote recall and simple calculation questions, and including more opportunities to write by synthesizing evidence to support claims.

SOCIAL STUDIES

Insight used criteria derived from leading assessment evaluation tools, and appropriate for the purposes of this Audit, to assess social studies assessments on a 3-point scale. Evaluation tools of origin are linked.

Five Social Studies Assessment Criteria and 3-point scale

A. Texts, including primary sources, are worth reading. Texts reflect the quality of writing that is produced by authorities in the social sciences (*from Criteria for Procuring and Evaluating High-Quality Assessments, CCSSO*).
B. Informational texts and tasks associated with them reflect demands of shifts (citing evidence, building knowledge, text complexity) and standards (*from Criteria for Procuring and Evaluating High-Quality Assessments, CCSSO*)
C. Assessments and tasks should include purposefully linked passages or questions that allude to other accessible historical texts or sources (*from Assessment Evaluation Tool, achievethecore.org*)
D. Variety of item types to accurately assess a standard. Students should have the opportunity to write in response to high-quality texts and primary sources (*from Criteria for Procuring and Evaluating High-Quality Assessments, CCSSO*)
E. Items should be designed to elicit direct, observable evidence of the degree to which a student can independently demonstrate the key understandings relative to the demands of the social studies standard(s) (*from EQuIP rubric, achieve.org*)

3	a significant majority or all aspects of the criterion
2	some aspects of the criterion
1	none or nearly none of the aspects of the criterion

For shorter artifacts, Insight adapted the EQuIP Task Review Rubrics, also used by ISBE. This rubric assesses standard alignment, attention to teaching strategies and literacy strategies, and implementation support. Insight rated these artifacts with the following descriptors:

E	Most criteria checked
E/I	Many criteria checked but could use minor improvements
R	Some criteria checked
N	Task not recommended for instruction

SOCIAL STUDIES

Ten assessments were rated on the social studies Assessment Criteria:

Criteria	Average score (between 1-3)
A. Texts, including primary sources, are worth reading. Texts reflect the quality of writing that is produced by authorities in the social sciences.	2
B. Informational texts and tasks associated with them reflect demands of shifts (citing evidence, building knowledge, text complexity) and standards.	1.6
C. Assessments and tasks should include purposefully linked passages or questions that allude to other accessible historical texts or sources.	Not rated
D. Variety of item types to accurately assess a standard. Students should have the opportunity to write in response to high-quality texts and primary sources.	1.7
E. Items should be designed to elicit direct, observable evidence of the degree to which a student can independently demonstrate the key understandings relative to the demands of the social studies standard(s).	1.8

Three artifacts were rated on the adapted EQuIP Task Review Rubric. One artifact was a graphic organizer and was not rated.

Descriptor	Number of Artifacts
E: Most criteria checked	0
E/I: Many criteria checked but could use minor improvements	0
R: Some criteria checked	1
N: Task not recommended for instruction	2

SOCIAL STUDIES

FINDINGS

1. Assessments scoring higher on the criteria featured texts that were worth reading, including some common interims with one nonfiction text and a shorter primary source.

2. Many assessments contained mostly or all selected response, with little to no opportunities for students to write in response to any text, which accounted for lower scores on criteria A and D.

3. Many assessments contained questions that were based only on content recall or memorization, which accounted for lower scores on criteria B and D.

4. US History common interims analyzed had lexile appropriate texts worth reading, associated selected response, and a writing prompt. There is more potential to include an additional shorter primary source, and to connect the writing prompt to a text within the assessment, so students can cite from it in their responses.

5. There is more potential to include more cross-cutting concepts and science and engineering practices in assessments and tasks by including fewer rote recall and simple calculation questions, and including more opportunities to write by synthesizing evidence to support claims.

6. Assessments were not rated on criteria C as not enough the submitted assessments warranted items that asked students to consider multiple texts or sources

HEALTH & WELLNESS

With the exception of two IH Health/Wellness common interims, Insight adapted the EQuIP Task Review Rubrics, also used by ISBE, to assess the remaining nine Health/Wellness artifacts, which were all from Driver's Education or IH Driver's Education. This rubric assesses standard alignment, attention to teaching strategies and literacy strategies, and implementation support. Insight rated these artifacts with the following descriptors:

Descriptor	Number of Artifacts
E: Most criteria checked	0
E/I: Many criteria checked but could use minor improvements	9
R: Some criteria checked	0
N: Task not recommended for instruction	0

HEALTH & WELLNESS

Two IH Health/Wellness assessments were rated on the five science Assessment Criteria:

Criteria	Average score (between 1-3)
F. Assesses state science standards to provide evidence about students' achievement in science. Assessment requires students to use some understanding of Disciplinary Core Ideas to successfully complete it, and includes Reading and Writing for Science and Technical standards.	3
G. Assessment requires students to use at least one Science and Engineering Practice to successfully complete the task.	3
H. Assessment requires students to identify and interpret evidence and engage in scientific reasoning as they make sense of phenomena and address problems.	3
I. There are varied task types requiring a range of analytical thinking and cognitive complexity.	3
J. Majority of assessment cannot be answered without information from tasks or items, nor can the majority of the assessment's items be answered successfully by using rote knowledge.	3

HEALTH & WELLNESS

FINDINGS

1. Driver's Education and IH Driver's Education artifacts all received E/I ratings. Most artifacts were short-answer or selected-response heavy, which may mirror the item variety present in the Illinois learner's permit exam. There may be opportunities to include more writing prompts, so that students can continue to hone their writing skills in DE and IH DE.

2. The two IH Health / Wellness common interim assessments contained various item types, including selected response and short answer, and opportunities to populate a table. Students were required to read graphs and perform mathematical calculations. Illinois PE standards were stated and were aligned to the scope and sequence for the class. It is noted, however, that all the common interims for the class contain the same standards.

WORLD LANGUAGES

For all eleven World Languages (Spanish I or IH Spanish I) artifacts, Insight adapted the EQuIP Task Review Rubric, also used by ISBE. It assesses standard alignment, attention to reading and writing, and attention to speaking and listening. Insight rated these artifacts with the following descriptors:

E	Most criteria checked
E/I	Many criteria checked but could use minor improvements
R	Some criteria checked
N	Task not recommended for instruction

Descriptor	Number of Artifacts
E: Most criteria checked	8
E/I: Many criteria checked but could use minor improvements	3
R: Some criteria checked	0
N: Task not recommended for instruction	0

WORLD LANGUAGES

FINDINGS

1. The common interims for both courses feature ACTFL standards on interpersonal and cultural communication. The English and Writing standards (at at least the 9-10 level) also feature, especially when students are asked to cite evidence from text. These assessments feature two texts, one dialogue and the other from another genre.

2. Formative assessments analyzed show alignment to standards, and items within the context of a story. Texts for some assessments provide opportunities to build knowledge while reading and practicing. Others show a variety of task types, including multiple-select Selected Response.

UNIT PLAN ANALYSIS

Sample unit plans across the mathematics, English, science, and social studies content areas, when available, were analyzed according to criteria derived from the analysis tools noted.

Sample unit plans in wellness and world languages contain general notes on standards alignment (when applicable or appropriate) and usability.

MATHEMATICS

	Rubric for Lessons & Units: Math (from EQuIP Rubric)	
	I: Alignment to the Depth of the CCSS. Does the unit target a set of grade-level Illinois State math standards? Are the SMPs central to the lessons identified, handled in a grade-level appropriate away, and connected to the content being addressed? Is there a balance of rigor (procedures and conceptual understanding) appropriate for the standards being taught?	**II: Key Shifts of the CCSS: Is the unit focused on the parts of it that constitute major work of the grade standards? Does content build on previous understandings, and are there opportunities for students to connect knowledge and skills vertically and horizontally? Is there opportunity for the appropriate balance of application, conceptual understanding, and procedural skill and fluency relative to the standards being taught?**
AP Calculus AB, Unit 6, The Definite Integral	Lists standards. Chapter 6 of book. SMPs are also listed. Unclear how the listing of standards and SMPs drives the user of this plan to integrate them in lessons. Rigor balance not explained in this unit nor referred to; will need to look at lesson-level documentation and/or instruction.	Yes. Definite integral is major part of AB. Coherence opportunities and rigor balance will need to be examined at the lesson level. Unit plan does not mention these things as drivers of the unit.
Integrated Math I, Unit 2: Linear Equations	Standards listed in unit plan do not list all the standards in the curriculum map for this unit. Standards on unit plan do represent a balance of conceptual understanding and procedural skill.	Insufficient information about evidence of key shifts in materials and instruction methods, instructional supports, and assessments (not sure which summative is connected to this unit, if a Common Interim or curriculum-embedded assessment).
IH Integrated Math I, Unit 5:	Standards as listed in unit would yield a balance of conceptual and procedural. There is a focus on priority	Focus and rigor assumed by list of standards. Across-grade level and vertical coherence unclear

Linear Systems and Piecewise Defined Functions	standards listed in Chapters 11-13. Unclear whether math practices are prioritized as connections (even ones stated in textbook materials) aren't explicitly highlighted in the unit plan.	as opportunities to highlight those connections aren't listed in plan.
Integrated Math I Instructional	Nothing specific to Instructional sections. Curriculum survey results only have one Instructional teacher responding (IM III Instructional). Instructional teachers are expected to modify the core course(s). Unit plans for those core courses contain key vocabulary, some general strategies and graphic organizers, but lack content-specific accommodations and modifications documented and used in the unit plans.	
Integrated Math III, Unit 4: Rational/Radical Equations	Standards listed in unit are priority standard in syllabus (REI.A.2) and other standards listed in curriculum map. RST standards also listed. Envision chapters noted. Unit plan standards, essential questions, and learning targets should be parsed out over the 4 weeks with base materials listed. Connection to SMPs not apparent/highlighted in unit plan, nor are balance of conceptual and procedural.	Focus on just some of the standards. References to coherence (prior knowledge, etc.) lacking. Connect application/concept to procedural (which RST standards apply where?).

MATHEMATICS (cont.)

	Rubric for Lessons & Units: Math (from EQuIP Rubric)	
	III: Instructional Supports: Is the unit plan easy to understand and use? Does the unit use and encourage precise use of mathematics, terminology, and academic language? Does it provide *all* students (with scaffolding if necessary) with multiple opportunities to engage with problems and tasks that stimulate mathematical thinking? Does it integrate appropriate supports for students who are EL, have disabilities, or operate well below grade level?	**IV: Assessment: Does the unit regularly assess whether students are mastering standards-based content and skills using varied modes of curriculum-embedded pre-, formative, summative, and self-assessments? Do the assessments elicit direct, observable evidence of the degree to which students can independently demonstrate the major targeted grade-level standards?**
AP Calculus AB, Unit 6, The Definite Integral	May not be easily usable to some teachers without clear recommended sequence of lessons, texts, and assessments.	
Integrated Math I, Unit 2: Linear Equations	There is no guidance about how to use the resources listed (most especially, the Envision 2019 textbook, only saying the unit aligns to Topic 2).	Insufficient information about evidence of key shifts in materials and instruction methods, instructional supports, and assessments (not sure which summative is connected to this unit, if a Common Interim or curriculum-embedded assessment).
IH Integrated Math I, Unit 5: Linear Systems and Piecewise Defined Functions	Instructional supports towards specific topics aren't listed; only global literacy strategies listed at beginning of unit.	No clear assessment(s) for the unit.
Integrated Math I Instructional	Nothing specific to Instructional sections. Curriculum survey results only have one Instructional teacher responding (IM III Instructional). Instructional teachers are expected to modify the core course(s). Unit plans for those core courses contain key vocabulary, some general strategies and graphic organizers, but lack content-specific accommodations and modifications documented and used in the unit plans.	
Integrated Math III, Unit 4: Rational/Radical Equations	Instructional supports listed globally but not specific.	Assessments listed, but formatives and overall summative assessment limited in scope to just priority standard(s).

ENGLISH

	Rubric for Lessons & Units: ELA (from EQuIP Rubric)	
	I: Alignment to the Depth of the CCSS. Does the unit target a set of grade-level Illinois State ELA/Literacy standards? Does it include a clear and explicity purpose for instruction, and feature texts that measure within the grade-level text complexity band and are of sufficient quality for the stated purpose?	**II: Key Shifts of the CCSS: Does the unit feature close reading, a focus on text-based evidence, writing from sources, and academic vocabulary? Do the unit's texts build disciplinary knowledge, increase in complexity, and represent a balance of literary and informational? Is there a balance of on-demand and process writing? Is there evidence of short, focused research project(s)?**
AP Language and Composition, Unit 7: Qualifying/ Counterarguments, Multiculturality	Mentions texts from *Bedford Reader*, unclear which ones and thus how complex. Clear focus on multiculturality. Targets set of standards from College Board Unit 7.	Unclear, resources mentioned but not specific. Lessons themselves may include close reading, text-based evidence, writing from sources, but those are unavailable.
English I, Unit 3: Short Story	RL.1, RL.2, RL.3, and writing informative.	Many texts, assessment types, and graphic and other organizers are mentioned. Need to go into the lesson and instructional level to see if close reading, text-based evidence, and writing from sources is a focus. Academic vocabulary mentioned at end of unit but how it fits in comes down to the lesson level.
English I EL	Nothing specific to EL. English I EL sent English Resources folder; resources are general ed and must still be accommodated for English Learners. Curriculum survey results indicate EL teachers must modify the core course themselves.	
IH Survey of Literature, Unit 5: Making a Difference	RL.6, RL.2, writing process, writing over extended time. There is a clear theme--making a difference. The texts are Lizzie Bright recommended for 10-12 years old. I Have a Dream--lexile 930 (5th grade) and themes 6-8th grade.	Academic vocabulary is listed. Unit mentions skills to find main idea with strategies to use to find text-based evidence. Balance of literary and ifnormational text. Need to get to lesson level to determine how much these shifts are apparent day-to-day.
English II, Unit 6: Cultural Differences	RL.1, RL.6, L.1, L.4 Writing argumentative. Thematic unit on cultural differences. Texts are of appropriate complexity and publishable quality.	Lacking specifics on exactly how listed templates, strategies, and assignments will lead to student mastery. May not be easily usable to some teachers without clear recommended sequence of lessons, texts, and assessments.
IH World Literature, Unit 7: Modern Age - Looking Up	Writing organization, RI.2, RL.2, RL.3. Topical unit on renaissance lit with themes on modernization. Recommends teachers "select a Shakespeare play."	Lacking specifics on exactly how listed templates, strategies, and assignments will lead to student mastery. May not be easily usable to some teachers without clear recommended sequence of lessons, texts, and assessments. No formatives listed, summative assessments listed as what exists in Schoology.

ENGLISH (cont.)

	Rubric for Lessons & Units: ELA (from EQuIP Rubric)	
	III: Instructional Supports: Is the unit plan easy to understand and use? Does the unit cultivate student interest in reading, writing, and speaking about the texts? Does it provide *all* students (with scaffolding if necessary) with multiple opportunities to engage with text of appropriate complexity for the grade level? Does it integrate appropriate supports for students who are EL, have disabilities, or read well below grade level?	**IV: Assessment: Does the unit regularly assess whether students are mastering standards-based content and skills? Do the assessments elicit direct, observable evidence of the degree to which students can independently demonstrate the major targeted grade-level standards with appropriately complex text?**
AP Language and Composition, Unit 7: Qualifying/ Counterarguments, Multiculturality	Need to see resources. Unclear how EL, students with disabilities, and students unable to access texts without supports will be supported.	Common Interims provide lengthy grade-level texts. Is the amount of writing compared to selected response proportional to actual AP test?
English I, Unit 3: Short Story	Many graphic and other organizers are mentioned. Separate standards for EL students re: certain texts, are linked.	Many assessments are mentioned. The order in which they're given, and why, should be explained.
English I EL	Nothing specific to EL. English I EL sent English Resources folder; resources are general ed and must still be accommodated for English Learners. Curriculum survey results indicate EL teachers must modify the core course themselves.	
IH Survey of Literature, Unit 5: Making a Difference	Instructional supports listed, but lacking specifics on exactly how listed templates and assignments will lead to student mastery. May not be easily usable to some teachers without clear recommended sequence of lessons, texts, and assessments.	Assessments are listed, unclear which order they should be administered.
English II, Unit 6: Cultural Differences	Lacking specifics on exactly how listed templates, strategies, and assignments will lead to student mastery. May not be easily usable to some teachers without clear recommended sequence of lessons, texts, and assessments.	
IH World Literature, Unit 7: Modern Age - Looking Up	Lacking specifics on exactly how listed templates, strategies, and assignments will lead to student mastery. May not be easily usable to some teachers without clear recommended sequence of lessons, texts, and assessments. No formatives listed, summative assessments listed as what exists in Schoology.	

SCIENCE

	EQuIP Rubric for Lessons & Units: Science Category I: NGSS Design		
	A: Does making sense of phenomena and/or designing solutions to a problem drive student learning in the unit?	**B: Does the unit build understanding of multiple grade-appropriate elements of the SEPs, DCIs, and CCCs by developing and using them?**	**C: Does the integration of CCCs, SEPs, and DCIs drive student performances?**
Biology, Unit 1: Metabolism	Somewhat. Description says students will explain and model to understand photosynthesis and respiration (supporting them in sense-making), but lacks explanation of how the essential questions tap into prior knowledge.	Somewhat. The DCI of Life Science is inherent in the standards, but references to SEPs and CCCs are oblique: one of the learning intentions is for students to construct a model to show how photosynthesis and cellular respiration use organic molecules to store energy. Labs are mentioned as formatives. However, no essential questions explicitly stated in the unit drive the teaching or selection of materials that are listed. What patterns should students notice? It is assumed that structure and function CCC will be addressed through creation of photosynthesis models.	Unclear. Some formatives list labs but no essential questions guide them.
IH Biology, Unit 4: Genetics	Verbs in the objectives are mostly explain, describe, identify, and summarize. These do not explicitly evidence sense-making in the ways that the essential questions listed in the unit do. How do the essential questions translate to the ways that students will make sense of the phenomena (by discussing, arguing, comparing/contrasting, relating, etc.)?	Assessments for all 4 sections mention labs and some written products. Models and structure and function are assumed when dealing with DNA and genetic processes. Arguing from evidence possible when discussing GMOs. Need more clarity about how the integration of these dimensions will drive the unit.	See B.
Chemistry, Unit 3: Periodic Trends	Students use structure of the PT to make sense of the PT's organization and effects on element behavior. Unit plan lists labs, CERs, model-building. Mentions relying on previous understandings	Graphing, recognizing patterns, use of math, applies the concepts of stability and change and energy and matter (per the activities listed).	Depending on content of labs, yes. Need more information about the nature of the summative questions guiding the creation of projects, the questions guiding the completion of labs, and

	(including graphing) to make sense.		the specific questions driving the CERs. Including these in the unit plan is helpful to read for someone who is backwards planning.
IH Chemistry, Unit 12: Solutions and Reactions in Solutions	No. Essential questions are "what" questions and there are no learning targets or objectives different from the academic standards.	Listed assessments include a Titration Lab and Report. Assumptions of measurement and analyzing and interpreting data. Other than that, no mention of how teaching and learning elsewhere will be driven by how concepts will cut across other concepts, and/or by how students will use science and engineering practices to understand the DCIs. POGILs listed as resources/activities but nothing listed in unit plan about what guides them.	Listed assessments include a Titration Lab and Report. Assumptions of measurement and analyzing and interpreting data.

SCIENCE (cont.)

	EQuIP Rubric for Lessons & Units: Science Category I: NGSS Design		
	D: Do the lessons fit together to target a set of performance expectations? Do they build on prior lessons?	**E: Are there opportunities, where appropriate, to link life, physical, and earth/space science?**	**F: Are there connections to Math and/or ELA?**
Biology, Unit 1: Metabolism	Unclear. Resources are listed but sequence of lessons and their intentions is unclear or up to the teacher.	Unclear. If they are embedded in materials or specific lessons, these linkages do not drive the overall direction of the unit.	Yes. Standards are listed in conjunction with the creation of lab reports.
IH Biology, Unit 4: Genetics	The four components of the unit have vocabulary words and essential questions that fit a sequence. However, the performance expectations seem to remain as a list of assessements/evidence that do not change from component to component.	Stays in Life Science.	ELA standards listed, and math inherent in study of allelle distribution.
Chemistry, Unit 3: Periodic Trends	Specific lessons aren't mentioned. The topics and sequence of questions cohere, and there are mentions of concepts building based on prior knowledge, but gradations aren't lesson by lesson. This trajectory would be helpful to understand how students are expected to build knowledge and when.	There are articles and videos. If linkages are discussed there, they can be mentioned in the unit plan so as to elevate those linkages to someone using this Unit Plan to highlight them in discussion or assessment.	ELA standards listed, and math assumed in labs.
IH Chemistry, Unit 12: Solutions and Reactions in Solutions	Standards are listed including the mass conservation standard, which requires knowledge of atoms. No mention of coherence elsewhere.	Unless mentioned in materials and lab, no evidence in unit plan mentioning these linkages specifically.	ELA standards listed, and math assumed if there is a titration lab.

SOCIAL STUDIES

	Rubric for Lessons & Units: Social Studies (from Washington Quality Review Rubric, adapted from EQuIP Rubrics)	
	A: Alignment to Standards: Does the unit target a set of grade-level standards in the Illinois Learning Standards in one or more of the following areas: Geography, Civics, Economics and Financial Literacy, History, or K-12 Inquiry Skills? Does it integrate social studies content knowledge with reading, writing, speaking, and listening skills as outlined in the ELA and Literacy in History/Social Studies Standards?	**B: Teaching Strategies: Does the unit integrate content and skills, contain inquiry-sparking questions, and encourage the 6Cs in 21st century learning? Does it incorporate good literacy practices with increasingly complex text?**
US History, Unit 7: The 1920s	History standards noted, R.HST standards assumed. Should it contain EC standards as well to account for inequality?	Texts noted through textbook and some linked resourced. Essential questions are inquiry-sparking--does the unit require that they be asked as noted in the assessments?
IH US History, Unit 2: American Revolution and Early Republic	History standards noted both primary and supporting, R.HST standards assumed.	Texts noted through textbook and some linked resourced. Essential questions are inquiry-sparking--does the unit require that they be asked as noted in the assessments?
IH Global Studies, PMSA Syllabus	IH Global Studies not located in Sharepoint; using only PMSA Syllabus for purposes of Audit: Textbook used is Ways of the World, 2nd ed., Strayer (2013).	Syllabus encourages students to use Historical Reasoning Skills (Contextualization, Comparison, Causation, Continuity and Change Over Time) as defined by the College Board. No units located in Sharepoint in order to examine integration of content and skills, inquiry-sparking essential questions, and incorporation of literacy strategies with complext text.

SOCIAL STUDIES (cont.)

	Rubric for Lessons & Units: Social Studies (from Washington Quality Review Rubric, adapted from EQuIP Rubrics)	
	C: Instructional Supports: Is the unit responsive to varied student learning needs (including but not limited to tech, engagement, prior learning, support with CER, etc.)?	**D: Assessment: Do all assessments in the unit elicit evidence that a student can independently demonstrate that they can meet the targeted standard(s) identified in it? Are there varied assessment types throughout the unit?**
US History, Unit 7: The 1920s	Links graphic organizers to support CER.	Lists different assessment types.
IH US History, Unit 2: American Revolution and Early Republic	Links graphic organizers to support CER, argumentative writing template.	Lists different assessment types.
IH Global Studies, PMSA Syllabus	No units located in Sharepoint to determine this.	Potential assignments in syllabus listed: interactive notebooks, class discussions, textbook-based assignments, academic and historical vocabulary, individual and group activities, digital resource projects, student use of technology, map exercises, major projects, writing exercises (including SAQs, DBQs, and Long Essay Qs), unit tests. No units or curriculum map located in Sharepoint to see cadence and frequency of assignments/assessments.

WELLNESS

Driver's Education	Unit 2: Being a Responsible Driver Standards and Resources are stated. Unit mentions IL's Rules of the Road book, from which requirements of this class must be based. Overall, unit should have a clear trajectory/arc explaining how the resources and texts will build within the unit towards culminating questions. Resources included are helpful, but sequencing is necessary to build a story, clear use of how essential questions will drive daily activities and discussions, and how final summative assessments will influence the formative assessments before it.
IH Driver's Education	Unit 1: Permit Information and the Driving Task Standards and Resources are stated. Unit mentions IL's Rules of the Road book, from which requirements of this class must be based. Overall, unit should have a clear trajectory/arc explaining how the resources and texts will build within the unit towards culminating questions. Resources included are helpful, but sequencing is necessary to build a story, clear use of how essential questions will drive daily activities and discussions, and how final summative assessments will influence the formative assessments before it.
Sophomore PE	Fitness Unit. Unit has IL state standards, national health standards, and writing standards cross-walked with one another. Additionally, RST and WHST standards are aligned below. Unit components and subheadings include component-specific vocabulary. Unit, unit should have a clear trajectory/arc explaining how the resources and activities will build within the unit towards some overall assessment marking the end of the unit. Activities included are helpful, but sequencing is necessary to ensure enough time is given to each component within the four weeks of the unit.
IH Health & Wellness II	Unit 1. Unit has IL state standards, national health standards, and writing standards cross-walked with one another. Additionally, RST and WHST standards are aligned below. Unit components and subheadings include component-specific vocabulary. Unit should have a clear trajectory/arc explaining how the resources and activities will build within the unit towards some overall assessment marking the end of the unit. Activities included are helpful, but sequencing is necessary to ensure enough time is given to each component within the four weeks of the unit.

WORLD LANGUAGES

Spanish I	Saludos (2A) Unit contains essential questions, essential vocab, resources, and literacy targets addressed in the unit. Unit should have a clear trajectory/arc explaining how the resources and texts will build within the unit towards culminating assessment(s). Resources included are helpful but sequencing is necessary to ensure explicit inclusion of prior knowledge, clear use of how essential questions will drive daily activities and discussions, and how final summative assessments will influence the formative assessments before it.
IH Spanish I	Mi Familia (3A) Unit contains essential questions, essential vocab, resources, and literacy targets addressed in the unit. Unit should have a clear trajectory/arc explaining how the resources and texts will build within the unit towards culminating assessment(s). Resources included are helpful but sequencing is necessary to ensure explicit inclusion of prior knowledge, clear use of how essential questions will drive daily activities and discussions, and how final summative assessments will influence the formative assessments before it.

CURRICULUM MAP / SCOPE AND SEQUENCE NOTES

The following are notes regarding Scopes and Sequences, and/or Curriculum Maps, for the analyzed courses.

MATHEMATICS

AP Calculus AB	Units 2-7 sequence generally follows the 8 2020 College Board units (with Unit 8 being review of year and Unit 1 being review of prior learning).
Integrated Math I	Missing standards: HSN-Q.A.2, HSA-SSE.A.1, HSA-REI.D.10, HSF-IF.B.6, HSF-IF.C.9, HSF-BF.A.1, HSF.LE.A.3, HSF.LE.B.5, HSS-ID.C8, HSS-ID.C9, HSG-CO.A.2, HSG-CO.A.3, HSG-CO.A.4, HSG-CO.A.5, HSG-CO.B.6, HSG-CO.B.7, HSG-CO.D.13, HSG-GPE.B.4, HSG-GPE.B.7. Standards on scope and sequence that are not addressed in IM I list of Standards: HSA-REI.B.4, HSN-RN.A.1, HSN-RN.A.2, HSA-SSE.B.3.C, HSF-BF.A.2, HSG-CO.C.9 (H), HSG-GPE.B.6, HSG-CO.C.10 Textbook in use is Savvas IM I: Green on both alignment and usability on EdReports.
IH Integrated Math I	Missing standards: HSF-IF.B.5 Standards on scope and sequence that are not addressed in IM I list of Standards: A-SSE.B.3c, F-IF.C.8b, G-CO.C.9, G-MG.A.1, G-MG.A.3, G-CO.C.10, G-CO.C.9, G-MG.A.3, G-SRT.B.5, G-CO.C.10, G-CO.C.11, G-SRT.B.5. Textbook in use is HMH Integrated Math I (2015). Red on alignment and not rated on usability on EdReports. CID is missing. All the common interims the same. Same standards, same structure, same kind of "create a system of equations" short answer/OR.
Integrated Math I Instructional	Nothing specific to Instructional sections. Curriculum survey results only have one Instructional teacher responding (IM III Instructional). Instructional teachers are expected to modify the core course(s).
Integrated Math III	Missing standards: HSS.IC.A.1, HSS.IC.A.2, HSS.IC.B.3, HSS.IC.B.4, HSS.IC.B.5, HSS.IC.B.6, HSS.MD.B.6, HSS.MD.B.7, HSN.CN.C.8, HSN.CN.C.9, HSA-SSE.B.4, HSA-APR.A.1, HSA-APR.B.2, HSA-APR.C.4, HSA-APR.C.5, HSA-REI.D.11, HSG-SRT-D.9, HSG-SRT-D.10, HSG-SRT-D.11, HSF-IF.B.4, HSF-IF.B.5, HSF-IF.C.8, HSF-IF.C.9, HSF-BF.A.1, HSF-BF.B.3, HSF-BF.B.4, HSF-LE.A.4, HSG-GMD.B.4, HSG-MG.A.1, HSG-MG.A.2, HSG-MG.A.3, HSA-CED.A.4

	Standards on scope and sequence that are not addressed in IM III list of Standards: HSA.SSE.B.3, HSA.CED.A.2, HSA.REI.D.10, HSS.ID.B.6, HSN.RN.A.2, HSS.ID.B.5, HSS.CP.A.1, HSS.CP.B.7, HSS.CP.B.8, HSS.CP.A.4, HSS.ID.A.3, HSG.GPE.B.4, HSN.Q.A.2, HSG.GPE.B.5, HSG.GPE.B.7, HSG.C.A.2, HSG.GPE.A.1, HSG.C.B.5, HSG.C.A.3, HSG.SRT.C.8, HSF.TF.A.3. Priority standards schedule for 8 units lists mostly priority standards for the IM III course, with the exception of Unit 6 and ID.B.5 in Unit 5. Syllabus mentions Envision textbook (assuming Integrated Math) and IXL. Unit plans list which Envision chapters align to units. Syllabus mentions summative assessments and grading policy for them based on mastery based learning. However, the summative assessments are each based on one priority standard. The priority standard is then subdivided into 8 categories, and continued mastery of this yield the grade for the entire unit. The priority standard should not constitute the entire summative and thus the grade (based on MBL) for the entire unit, as there are other standards.

ENGLISH

AP Language and Composition	follows AP scope and sequence Is the AP Lang and Composition test required testing for students in that class? If so, does the grade figure in to the class as a final exam?
English I	focus on R.1-2 throughout, attached texts show commitment to reading complex texts if L1, 2, and 4, include W4-6 as focus; where are R.5-6 (evidence of author's purpose/POV and how author uses rhetoric/language to develop it)?; where is R/W standard 9—linking texts and alluding to other texts? Holt McDougal Literature is partially meets for 9-10, does not meet for 11-12, according to EdReports.com Each unit requires texts from different sources. Teachers and staff report resources are inconsistently accessible, leading to (or coincidental with) inconsistent expectations about minimum usage. Holt McDougal Literature textbook either partially or does not meet expectations, according to EdReports. A high-quality, all-inclusive textbook will provide the resources needed that will ensure grade level standards are taught with enough coverage and review because texts and associated prompts and projects are designed with the skills and practices of the grade level in mind, building over course of the entire year.
English I EL	Nothing specific to EL. English I EL sent English Resources folder; resources are general ed and must still be accommodated for English Learners. Curriculum survey results indicate EL teachers must modify the core course themselves.
IH Survey of Literature	RL.1 not listed in scope but likely should be (citing evidence in fictional texts). RI/HST.1 and W.1 should also be listed in scope. where is R/W standard 9—linking texts and alluding to other texts? L.2 not listed nor SL.2, .3, .5, .6.
English II	focus on R.1-2 throughout, attached texts show commitment to reading complex texts if L1, 2, and 4, include W4-6 as focus; where are R.5-6 (evidence of author's purpose/POV and how author uses rhetoric/language to develop it)?; where is R/W standard 9—linking texts and alluding to other texts? Holt McDougal Literature is partially meets for 9-10, does not meet for 11-12, according to EdReports.com Each unit requires texts from different sources. Teachers and staff report resources are inconsistently accessible, leading to (or coincidental with) inconsistent expectations about minimum usage.

	Holt McDougal Literature textbook either partially or does not meet expectations, according to EdReports. A high-quality, all-inclusive textbook will provide the resources needed that will ensure grade level standards are taught with enough coverage and review because texts and associated prompts and projects are designed with the skills and practices of the grade level in mind, building over course of the entire year.
IH World Literature	RL.1 not listed in scope but likely should be (citing evidence in fictional texts). where is R/W standard 9—linking texts and alluding to other texts? L.2 not listed nor SL.2, .3, .5, .6.

SCIENCE

Biology	Both curriculum map and scope and sequence both essential provide the same information: standards associated with each of the five units. The curriculum map provides resources links, and the S and S show the length of time recommended for each standard. Since some lessons will integrate many standards, scope and sequence should reconsider blocking amounts of time to only one standard. If scope and sequence would more clearly list which materials will drive lessons, and those materials are aligned with the NGSS 3D Design approach, then the unit can have a base of standards integration from which to iterate if the teacher wishes.
IH Biology	Curriculum Map includes 8 units, standards per unit, and target timeframes. Resources are not listed. In unit plans, resources are listed but not in any particular order. Additionally, unit plans do mention essential questions, but end of unit assessments that should drive the unit are not explicitly noted. This information should appear in at least one of the documents (curriculum map/scope and sequence or unit plan(s)). Additionally, all of the common interims mainly assess one SEP and have the same cadence of questions (reading text, determining independent or dependent variable). The common assessments individually are strong. Do they drive the content and approach of the class?
Chemistry	Curriculum Map includes 8 units, with standards and SEPs per unit + target timeframes. Many resources listed. Scope and sequence includes timeframes for each unit plus whether standards are SOS.
IH Chemistry	Curriculum Map includes 15 units, standards per unit, and target timeframes. Resources are not listed. In unit plans, resources are listed but not in any particular order. Additionally, unit plans do mention essential questions, but end of unit assessments that should drive the unit are not explicitly noted. This information should appear in at least one of the documents (curriculum map/scope and sequence or unit plan(s)).

SOCIAL STUDIES

US History	Curriculum Map includes 13 units, with standards per unit + target timeframes. Textbook resource is only one listed. Scope and sequence includes timeframes for each unit plus whether standards are SOS. In unit plans, resources are listed but not in any particular order. Additionally, unit plans do mention essential questions, but end of unit assessments that should drive the unit are not explicitly noted. This information should appear in at least one of the documents (curriculum map/scope and sequence or unit plan(s)).
IH US History	Curriculum Map includes 13 units, with standards per unit + target timeframes. Textbook resource is only one listed. Scope and sequence includes timeframes for each unit plus whether standards are SOS. In unit plans, resources are listed but not in any particular order. Additionally, unit plans do mention essential questions, but end of unit assessments that should drive the unit are not explicitly noted. This information should appear in at least one of the documents (curriculum map/scope and sequence or unit plan(s)). This document is the same as the non-IH US History course above. Curriculum Handbook says IH Courses "are designed to further prepare students for more rigorous academic work and assessments. IH courses focus on in-depth exploration of themes and topics in the various academic areas, are student-centered, and include project-based experiential learning components. Additionally, these courses are developed with an emphasis on global awareness and international mindedness." Scope and sequence and curriculum map should reflect these additions.
IH Global Studies, PMSA Syllabus	IH Global Studies not located in Sharepoint; using only PMSA Syllabus for purposes of Audit. Textbook used is Ways of the World, 2nd ed., Strayer (2013).

WELLNESS

Driver's Education	Curriculum Map and Scope and sequence contain three units. Standards listed along with Reading and Writing standards. Standards from American Driver and Traffic Safety Education Association Novice Driver Education Curriculum Standards. Does the map and S&S meet this requirement from ilsos.gov: Section 27-24.2 mandates that all public schools must include the content of Chapters 11, 12, 13, 15, and 16 of the Illinois Vehicle Code, the administrative rules adopted pursuant to those chapters insofar as they pertain to the operation of motor vehicles, and the portions of the Litter Control Act [415 ILCS 105/1 et seq.] relating to the operation of motor vehicles. The Illinois Secretary of State produces the publication The Rules of the Road to inform potential drivers of all pertinent laws in Illinois.
IH Driver's Education	Curriculum Map and Scope and sequence contain three units. Standards listed along with Reading and Writing standards. Standards from American Driver and Traffic Safety Education Association Novice Driver Education Curriculum Standards. Does the map and S&S meet this requirement from ilsos.gov: Section 27-24.2 mandates that all public schools must include the content of Chapters 11, 12, 13, 15, and 16 of the Illinois Vehicle Code, the administrative rules adopted pursuant to those chapters insofar as they pertain to the operation of motor vehicles, and the portions of the Litter Control Act [415 ILCS 105/1 et seq.] relating to the operation of motor vehicles. The Illinois Secretary of State produces the publication The Rules of the Road to inform potential drivers of all pertinent laws in Illinois.
Sophomore PE	Curriculum map with 5 units. Scope and sequence empty but checklist provides crosswalk of unit characteristics, Illinois State PE standards 19-24, and RST and WHST standards.
IH Health & Wellness II	Curriculum map with 7 units. Scope and sequence has standards indicated SOS. Illinois State PE standards 19-21, along with RST and WHST standards. Differs from Sophomore PE in that it spends more time on fitness monitoring, discussion on healthy living and lifestyles, and fitness goals. In terms of common interims, CIA, CIB, and CIC are the same.

WORLD LANGUAGES

Spanish I	Curriculum Map has 8 units, 4 units divided into A/B. Scope and sequence contains SOS standards, chapters in textbook, and ACTFL standards alongside ELA Reading and Writing standards. Timeframes included in S&S.
IH Spanish I	Curriculum Map has 8 units, 4 units divided into A/B. Scope and sequence contains SOS standards, chapters in textbook, and ACTFL standards alongside ELA Reading and Writing standards. Timeframes included in S&S. Materials here seem identical to Spanish I. Curriculum Handbook says IH Courses "are designed to further prepare students for more rigorous academic work and assessments. IH courses focus on in-depth exploration of themes and topics in the various academic areas, are student-centered, and include project-based experiential learning components. Additionally, these courses are developed with an emphasis on global awareness and international mindedness." Scope and sequence and curriculum map should reflect these additions.

Gap Analysis and Action Plan for Priority Improvement Initiatives

The Gap Analysis phase aims to understand the barriers between the current conditions and the ideal outcomes.

This phase consisted of twelve hours of interviews with twelve different groups of stakeholders. Insight analyzed the contributions of participants for common trends that would yield priority actions the district should undertake to improve curriculum.

Finally, the data and evidence gathered from Curriculum Review and Gap Analysis were used to generate the ten suggested next steps in this report.

We hope the findings and recommendations in this report can serve as a springboard to the planning phase. We are so grateful for the opportunity to partner with you and look forward to supporting you on this journey in service of students.

FOCUS GROUPS

To illuminate both currently successful practices to be replicated, as well as identify pain points, twelve groups of stakeholders were invited to participate in hour-long focus group conversations. The focus groups were:

- math teachers
- social studies teachers
- English teachers
- science teachers
- health/wellness teachers
- world languages teachers
- district coordinators (3 groups)
- principals
- grade level principals
- students

Findings from these focus groups will be included in the Recommended Next Steps section of this report.

Insight thanks all participants for their time and valuable input during the process.

RECOMMENDED NEXT STEPS

1 Streamline Course Catalog by Prioritizing Core Graduation Requirements

It is recommended that the district create more space in school schedules for core courses needed for graduation, especially courses at the freshman and sophomore level. The district should audit the course catalog for opportunities to condense or eliminate specialty classes in favor of these core courses. Essentially, having fewer courses concentrates the district's curriculum improvement efforts and allows greater visibility into, and support for, the core courses necessary for progression through the high school grades leading to graduates ready to tackle 21st century college and career.

More sections of required courses concentrates the district's resources and staff towards the foundations required for student success. According to a focus group participant, "the underlying thing is equity. We have so many classes that start with 35 students, and in a specialty class, we have 11. In December, we figured out that a class had zero students. So that teacher had zero students the whole semester for that period." The district should examine whether district funds and space could be allocated more towards core courses; mis-allocation away from core freshman courses should not contribute to the currently 74% on-track rate of freshmen.

Concentrating the district's resources on core courses decreases the number of students per section. Teams of teachers with more consistent course loads are able to collaborate more on similar content, and create more opportunity to collaboratively problem-solve or provide interventions to groups of students (vs. planning or teaching other preps).

One focus group participant noted an additional potential benefit to diverse learners throughout the district: fewer courses for which to provide modifications and accommodations means more concentrated ability to provide services to students who require them, as well as plan the course's direction considering the needs of all students in co-taught classes.

This change would not require hiring new teachers, but rather, a realignment of the courses taught by current Proviso staff. Additionally, a smaller number of courses eases the burden on school leaders and school counselors who must ensure that prerequisites are met and class sizes remain manageable.

The district should also explore courses that can be combined with currently existing courses, in order to become a unit or section within that course.

RECOMMENDED NEXT STEPS

2 Increase School-Based Instructional Support for Teachers

All content area Focus Groups, and most coordinator and school leaders groups, mentioned the need for school-based content leads. This person may have both teaching and content lead duties.

Groups of teachers criticized the lack of a liaison between teacher groups and school and district leadership, as well as the lack of a figurehead with both content expertise and decision-making latitude (not to mention, time in their schedules to provide that support to their colleagues). One participant noted, "I have nowhere to go" for help with content. Another noted that, in the absence of a figurehead, teachers look like they're "going rogue," but the reality is that they do not have direction or guidance. "A room of equals is not the way to move things forward," noted one participant.

During the occasion that teachers of the same course or content have common planning time, the need for a content lead to lead conversations about data, push forward conversations about operations, and generally push forward school or district priorities around professional development, is apparent. Without it, one participant noted, "we're just going to be doing our own individual work."

Content leads may provide direct answers and swifter communication, which many participants noted was a problem. One participant noted that there is "no reliable chain of command" and without this, "there is no follow-through." Several participants noted the lack of follow-through on various things like materials availability and communication created an erosion of trust in the entire system.

The district should consider supporting content leads across the campuses, in order to provide teachers the leadership and support they feel would strengthen academic culture at PTHS D209.

RECOMMENDED NEXT STEPS

3 Prioritize Space in School Schedule for Collaborative Planning Time

As in the previous Next Step, all participants noted the need for more structured collaboration time among teachers. It is recommended that the district have PLC structures that establish clear student-centered outcomes, meet regularly, and whose products are connected to school- and district-wide goals.

As Focus Groups discussed materials consistency and availability, it became clear that lack of collaboration time among teachers may lead to less curriculum consistency overall. Additionally, some participants felt like they were missing opportunities to strengthen the overall curriculum by not collaborating. Several groups noted the weaker vertical alignment among courses in the same department, because there is no time to compare curriculum maps (and other long-term planning documents) with members of their own departments. "If we don't get together with each other, and really work on curriculum, we don't have curriculum," one participant said. Another participant noted, "our curriculum 9-12 has been piecemeal in the department, because we created it over years and there's been no time to sit and look at it."

Without this time, participants largely report feeling on their own. "I often can't find anything, and I'm lucky my colleagues help me and are able to share resources with me. Because otherwise, there's nothing. We don't have time to sit together and talk about curriculum, because we're just so busy and running after stuff."

For co-taught courses, participants report a deficit in time for co-teachers to plan with their general education counterparts. If they "don't have time to sit together and plan over what will be presented to students, students don't receive modified curriculum."

Some groups mentioned wanting more time to collaborate across schools. One participant said, "we need to sit down as a group and get clear on what the classes for our content look like. We just have no time to do that. I think we are dedicated teachers who would be committed to doing that, but there's no time during the day or no opportunities to do so." Blocks of time during Professional Development days was suggested as an opportunity for cross-school collaboration.

Participants noted the helpfulness of past opportunities to engage in summer planning. Stipends had been offered in the past; many participants welcomed the idea of spending some days between school years vetting and modifying high-quality curricular resources (including assessments) rather than spending time writing assessments.

RECOMMENDED NEXT STEPS

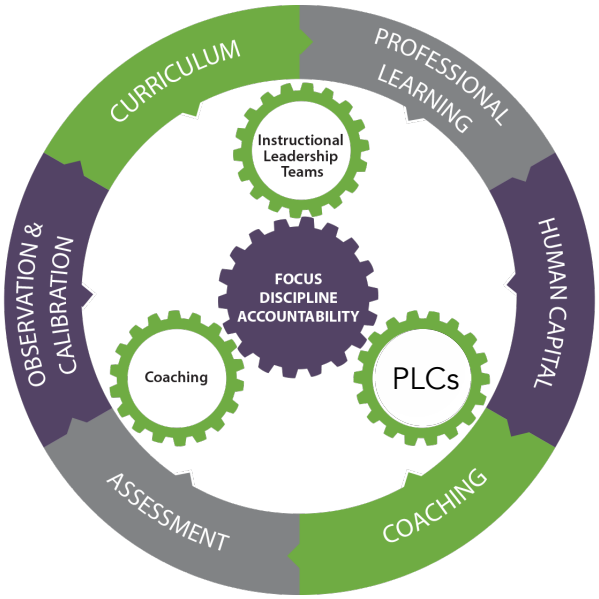
4 Implement Instructional Leadership Team Structure

Based on the Focus Groups, Insight determined there is an opportunity at each school to have Instructional Leadership teams consisting of the principal, GLPs, and representatives from different contents and grade levels.

The sample model below shows how leadership teams identifying high-lever initiatives can lead to student success and execute those initiatives with focus, discipline and accountability, as Instructional Leadership Teams (ILTs) "move the gears" of coaching and PLCs to drive forward initiatives. By instituting Instructional Leadership Teams, campus leadership teams can remain focused, make sure all initiatives are on track, solve issues, and communicate in order to gain traction toward attainment of goals.

Some participants are experimenting with this structure already, and report benefits of focus, shared vision, and accountability--not just with curriculum-related matters. "We work on PLCs and developing common assessments. Then we introduce our work to everyone at staff meeting, and then during the week during planning periods. Department or content area teams work to carry those action items out to their groups."

The limiting factor remains time, as Instructional Leadership Team (ILT) members must meet before or after school.



RECOMMENDED NEXT STEPS

5 Establish Curriculum Adoption Steering Committee to Create Schedule for Materials Adoptions and Associated Professional Learning

One of the best ways to dramatically improve student learning and engagement is to give teachers access to high-quality instructional materials (HQIM) and the support they need to use those resources (Chiefs for Change, 2018). Curriculum surveys, focus groups, and curriculum review of unit plans and artifacts reveal the need for consistent, high-quality instructional materials to be adopted and made available across the district.

All focus groups expressed a desire for more consistency of course scope, as well as availability and use of high-quality instructional materials.

Teachers noted the need for teacher "choice and voice" in the matter: teachers need to be the most critical voice in the conversation when it comes to choosing materials that would be chosen as base materials for courses. "I think if we were able to give our opinions and look into it ourselves, there's a better chance it will be received positively," one participant said.

Insight recommends establishing a district curriculum adoption steering committee that will determine the schedule for materials adoptions, as well as determine the quality and scope of professional development to help launch, guide and successfully implement any new curriculum.

Additionally, smaller content curriculum teams should be established to vet materials, achieve district-wide consensus and communication, and eventually be spokespeople for the adoption and successful implementation of the curriculum at their schools. Focus group participants noted the need to begin the focus on ninth grade core courses, which reflects the first Recommended Next Step in this report.

One participant notes, "there's way too much well-researched, reliable material out there. I don't want us to create the wheel again." Many participants remarked that purchased, high-quality curriculum materials are a "time-saver," and a "way to work smarter, not harder."

The district should note that, in addition to securing teacher buy-in, curriculum implementation requires the effort of also investing in professional development to support it. If curriculum materials adoption is a possibility for the start of '23-'24, the district should consider assembling steering committees as soon as possible.

RECOMMENDED NEXT STEPS

6 Determine District-Wide Consistency in Unit Plan Content and Format

Many participants noted that unit plans in their current format do not provide the needed guidance for sustained forward planning and cohesion across content or grade level.

Curriculum Review of unit plans also revealed the opportunity for many plans to have more concrete information about materials to use, and sequence to follow, alongside the usually stated time frames.

The unit and lesson plan analysis tools used noted the importance of usability. Many participants remembered their experiences as first-year teachers, who either were, or would have been, helped by unit plan formats that contained information that saved them from spending the energy and time creating. Currently, there are teachers at Proviso who report starting the beginning of the year (or mid-year) and having to do exactly that--improvise, create, or as one participant noted, "go back to something familiar," because there is not enough guidance in the Unit Plan documents.

The district should build on the countless hours of work already poured in to the current unit plans. Many plans contain positive aspects as noted in the Curriculum Review section of this report. Insight recommends the district establish a team of inter-content members to determine the level of granularity that would render the unit plans more usable to all (especially new teachers) and aligned to expectations reflected in the analysis tools in this report, allowing for the kind of autonomy that many participants said they and their colleagues would like to retain.

RECOMMENDED NEXT STEPS

7 Concentrate Professional Learning on Strengthening Capacity to Teacher College- and Career-Ready Standards

Student focus group participants noted that the most challenging, rewarding, and fulfilling classes they attend "reflect college material," "improve my knowledge and problem-solving skills," and "help me think on my feet." Conversely, students in the group criticize classes in which they do "busy work," "do assignment after assignment of just drilling us" and "memorizing," and classes in which "the teachers talk the whole time or where we just watch a video." Participants noted that they wanted their classes at Proviso to resemble college-level activities: they want to "do more debates," but also "go home and study just like in college."

Curriculum review revealed many artifacts that were memorization and recall-type assignments. The Illinois Learning Standards encourage science and engineering practices, argumentation and citing textual evidence, synthesizing primary sources and other texts, and other 21st century learning skills that Proviso's students need in order to be competitive. This leads to the recommendation that Proviso invest direction, time, and effort on providing sustained professional development on strengthening teachers' capacity to teach to college- and career-ready standards. This is the very essence of Goals 2 and 4 (Academic Achievement and Talent Development) in the district's strategic plan.

Participants in focus groups stated their, and their colleagues', desire for professional development would push their teaching to incorporate these skills. For example, some participants noted that they had yet to receive any training on the NGSS (New Generation Science Standards), and that any training they had received was done of their own volition.

Others noted that current curriculum documents show an obsolete focus on knowing facts and content for content's sake, rather than seeing the curriculum as an opportunity to build the *skills* reflected in the ELA and Reading and Writing for Science and History standards. Doing the work of revamping unit plans, as well as considering adoption of high-quality curriculum materials, requires that teachers be supported to implement those plans and materials with quality professional development along the way. Leadership at schools charged with supporting those teachers to build, strengthen, and understand their curriculum materials to reflect college- and career-ready standards should have a baseline understanding of these expectations as a way to strengthen their own instructional leadership.

RECOMMENDED NEXT STEPS

8 Account For, and Organize, All Physical and Digital Resources

As previously noted in survey data, there is a need for the district to account for and organize all physical and digital resources. Basically: what is out there? Who has it, or where is it? And what is the system for distributing it, buying it, and keeping tabs on it?

Many participants note that the physical resources that do exist are obsolete, and no longer fit for teaching the content in the class (e.g., outdated textbooks). Others note their worry that courses are enrolled with students with the knowledge that curriculum materials for those courses either do not exist or will arrive too late (necessitating teacher creation or procuring of other curriculum materials until the arrival of previously promised materials). Additionally, participants note that there are products of long work sessions that may be physically stored somewhere, or whose digital files are missing or moved.

Finally, participants noted that the absence of consistency when it comes to curriculum materials logistics has a more nefarious effect: the erosion of trust in the Proviso system as a whole. One participant noted, "it's about follow-through. I want direct answers. I want Yes, or No--are we getting this book or not?" Others noted the stress of not having or not knowing where materials are has ripple effects on preparation and investment. "If I know where things are before the end of the year, I can prepare over the summer, prepare myself, find all the things I need. The most basic needs. If not, then you're not meeting the hierarchy of needs."

Before, or concurrent, with the establishment of Curriculum Adoption Steering Committees, it is recommended that Proviso establish similar working groups to embark on an audit of physical and digital materials, starting with the courses whose curriculum materials are highest priority to revamp. If teacher teams or other content teams across the district are willing to assist in these efforts, Proviso should consider enlisting their help.

Appendix

I. Detailed Artifact Analysis

II. Unit and Task Analysis Documents and Links to EQuIP, CCSSO, and AET Rubrics

DETAILED ARTIFACT ANALYSIS

As mentioned previously, the Curriculum Management Audit focused on a limited number of courses throughout the district. Teachers were randomly selected to contribute to the overall curriculum picture at PTHS D209 by providing artifacts for one of the courses in the Audit.

The courses, per school, were:

	Math	English	Science	Social Studies	Health / Wellness	World Languages
Proviso East	AP Calculus AB, Integrated Math I (2 classrooms), Integrated Math III	AP English Language and Composition, English I (2 classrooms), English II	Biology (2 classrooms), Chemistry (2 classrooms)	US History (2 classrooms), IH Global Studies	Driver's Education, Sophomore PE	Spanish I
Proviso West	Integrated Math I (2 classrooms), Integrated Math III	AP English Language and Composition, English I (2 classrooms), English II	Biology (2 classrooms), Chemistry (2 classrooms)	US History (2 classrooms), IH Global Studies	Driver's Education, Sophomore PE	Spanish I
PMSA	AP Calculus AB, IH Integrated Math I (2 classrooms)	AP English Language and Composition, IH Survey of Literature (2 classrooms), IH World Literature	IH Chemistry, IH Biology	IH US History, IH Global Studies	IH Driver's Ed, IH Health & Wellness II	IH Spanish I

MATHEMATICS

Insight used criteria derived from leading assessment evaluation tools and appropriate for the purposes of this Audit, to assess math assessments on a 3-point scale. Evaluation tools of origin are linked.

Five Math Assessment Criteria and 3-point scale

A. The set of items is clearly consistent with the most important content of the identified standard, and items should be designed to elicit direct, observable evidence of a student's ability to independently demonstrate competency (*from EQuIP Rubric, achieve.org*)
B. Item set is consistent with the standards' primary aspect of rigor (conceptual, procedural, and/or application) (*from Criteria for Procuring and Evaluating High-Quality Assessments, CCSSO*)
C. Assessments should contain a variety of item types (*from Criteria for Procuring and Evaluating High-Quality Assessments, CCSSO*)
D. Assessment should demonstrate authentic connections between the content standards and the eight Standards for Mathematical Practices (*from Assessment Evaluation Tool, achievethecore.org*)
E. Majority of items on the assessment come from major work of the grade (priority standards) (*from Assessment Evaluation Tool, achievethecore.org*)

3	a significant majority or all aspects of the criterion
2	some aspects of the criterion
1	none or nearly none of the aspects of the criterion

For shorter artifacts, Insight used the EQuIP Task Review Rubric for Mathematics, also used by ISBE. It assesses alignment to standards, attention to the instructional shifts in the standards, and implementation support. Insight rated these artifacts with the following descriptors:

E	Most criteria checked
E/I	Many criteria checked but could use minor improvements
R	Some criteria checked
N	Task not recommended for instruction

MATHEMATICS

The following tables show artifact numbers (assigned by the order in which they were analyzed) and their 1-3 ratings on Criteria A-F on the math assessment criteria, or E-N ratings on the adapted EQuIP task rubric.

	Artifact number									
Criterion	1	2	3	4	5	6	7	8	64	65
A	3	3	2	2	3	3	2	3	2	2
B	3	3	3	2	2	2	2	2	2	2
C	2	2	1	2	2	2	1	2	3	3
D	2	2	1	2	2	2	2	2	2	1
E	3	2	2	3	2	2	2	1	3	3

Criterion	66	67	68	69	70	71	84	92	93	94
A	2	2	3	3	E	E	1	E	E	E/I
B	2	2	3	3			1			
C	1	1	2	2			2			
D	1	1	3	3			1			
E	1	1	1	1			1			

Criterion	96	97	99	107	110
A			1		2
B			3		3
C			2		2
D			1		3
E	R	R	1	R	2

MATHEMATICS

The following short descriptions, by artifact number, provide additional information on the artifacts above.

1	IM 1 unit assessment. Envision/Savvas. Good variety. Standards appropriate. Needs open response.
2	IM 1. See #1. Includes 7th and 8th grade standards.
3	IM 1 unit assessment. Only one standard: is this quiz or test? Tests strategies vs. allowing students to choose. Little item variety.
4	IM 1 unit assessment. #5-10 do not ask why rational or irrational. Needs open response. Sharepoint
5	IM 1 Instructional course unit assessment. From Kuta. Needs some Selected Response.
6	IM 1 Instructional course unit assessment. From Kuta. Has a scaffolded open response. Needs some Selected Response.
7	IM 1 Final. No item type variety. From Quizizz. No congruence/geometry. Should there be (is this semester or year final)?
8	IM 1 Final (Instructional). Good item variety, needs Open Response. All 8th grade or REI, SSE, CED. Should there be fewer of these and more of congruence/geometry, functions, SP, etc (is this semester or year final)?
64	AP Calc AB CIA. 48% SR, 52% Open Response (50/50 on AP test). (need rubric references to AP mathematical practices and percentage grading breakdown for Selected Response and Open Response).
65	AP Calc AB CIE. 48% SR, 52% Open Response (50/50 on AP test). need rubric references to AP mathematical practices and percentage grading breakdown for Selected Response and Open Response). Also, by CIE, exam should be closer to length of AP exam, with more than 1 Open Response
66	IH IM 1 CIC. 2 ID'ed standards, 7-8 Selected Response items each. Mostly procedural. Mostly Selected Response and no short answer. Long form problem should be a short answer. Mostly procedural not many attempts to use SMPs. Only two ID'ed standards and mostly Selected Response, so short on major work of the grade for a common interim. Other common interims are the same standards.
67	IH IM 1 CIA. 2 ID'ed standards, 7-8 Selected Response items each. Mostly procedural. Mostly Selected Response and no short answer. Long form problem should be a short answer. Mostly procedural not many attempts to use SMPs. Only two ID'ed standards and mostly Selected Response, so

	short on major work of the grade for a common interim. Other common interims are the same standards.
68	IM 3 Standard 6, summative assessment 1 and 2. HSE-GPE.B.4 is an IM 1 standard. The entire exam is built from it. The summative assessment is only on one standard. Exam is lengthy, contains item variety (except Open Response) and has balance of rigor. However, it is a summative assessment and only contains mostly one standard.
69	IM 3 Standard 2, summative assessment 1 and 2. HSE-CED.A.2 is an IM 3 standard. The entire exam is built from it. The summative assessment is only on one priority standard. Exam is lengthy, contains item variety (except Open Response) and has balance of rigor. However, it is a summative assessment and only contains mostly one standard.
70	AP Calc formative. Study guide. Aligned to standards. Students apply procedural knowledge on concavity, increase/decrease, rel min/max, but also conceptual understanding to explain why some rel min/max don't exist. Task addresses critical content of differentiation. Questions written to pinpoint student understandings. Task asks students to analyze one function's attributes.
71	IH IM I formative. Systems of Equations substitution exploration. Aligned to standards. Students explore by first using prerequisite knowledge with substitution with single variable equation, students asked to explain why, explain their rationale, check for reasonableness, package procedure in their own words, then solve. Comprehensive, and on two pages.
84	IM 3 formative 1 standard 1. A-SSE.A.2. Does not assess the identified standards. All matching, substitution, and expression simplification, rather than rewriting expressions by using its structure. No open response. All procedural or identification.
92	IM 1, formative. HW from Envision. Aligned to standards, students use SMPs of looking for structure and making use of it (slope/intercept). Task has balance of rigor (application and procedural), and requires they integrate both. Items connect to observable parts of the standard. Accommodations for all learners should be examined at the lesson level.
93	IM 1, formative. HW from Envision. Aligned to standards, students use SMPs of looking for structure and making use of it (slope/intercept). Task has balance of rigor (application and procedural), and requires they integrate both. Items connect to observable parts of the standard. Accommodations for all learners should be examined at the lesson level.
94	IM 1, formative, quiz. HW from Envision. Aligned to standards, students use SMPs of looking for structure and making use of it (slope/intercept). Task

	has mostly procedural and needs short answer/open response. Has a multi-select SR. Items connect to observable parts of the standard. Accommodations for all learners should be examined at the lesson level.
96	IM 1 Instructional, formative. Students need to use understanding of solving equations, no SMPs. Procedural only. Task can be used to support a critical grade level process. As this is instructional course, what are the supports included to give students opportunity to use SMPs, write, explain, etc (examine this at the lesson level)?
97	IM 1 Instructional, formative. Students need to use understanding of solving equations, no SMPs. Procedural only. Task can be used to support a critical grade level process. As this is instructional course, what are the supports included to give students opportunity to use SMPs, write, explain, etc (examine this at the lesson level)?
99	IM3 formative. Items aren't consistent with the standard, much of it is middle school level. Item set is consistent in that it's mostly procedural. There is a variety of item types, but no Open Response. No SMPs. Not major work of IM3.
107	IM I Instructional, formative. Students need to use understanding of solving equations, no SMPs. Procedural only. Task can be used to support a critical grade level process. As this is instructional course, what are the supports included to give students opportunity to use SMPs, write, explain, etc. (examine this at the lesson level)?
110	IM 3 formative 2 for standard 6. Long formative, exam-length. Half is content from standard (or other GPE standards), half is middle school. Asks students to show evidence of independent competency by showing work. Item set consistent with primary aspect of procedural and conceptual. Variety of item types but no Open Response. SMPs used when proving and making use of structures. Half comes from major work, other half is prerequisites.

ENGLISH

Insight used criteria derived from leading assessment evaluation tools and appropriate for the purposes of this Audit, to assess English assessments on a 3-point scale. Evaluation tools of origin are linked.

Five English Assessment Criteria and 3-point scale

A. Texts are worth reading. Should include high quality texts at appropriate lexile and qualitative complexity (*from Criteria for Procuring and Evaluating High-Quality Assessments, CCSSO*)
B. Balance of genres when appropriate; reflect demands of shifts (citing evidence, building knowledge, text complexity) and standards (*from Criteria for Procuring and Evaluating High-Quality Assessments, CCSSO*)
C. Assessments and tasks should include purposefully linked passages or questions that allude to other accessible texts or passages according to RI/RL.9 (*from Assessment Evaluation Tool, achievethecore.org*)
D. Variety of item types to accurately assess a standard. Students should have the opportunity to write in response to high-quality texts (*from Criteria for Procuring and Evaluating High-Quality Assessments, CCSSO*)
E. Items should be designed to elicit direct, observable evidence of the degree to which a student can independently demonstrate the key understandings of the text relative to the demands of the standard (*from EQuIP rubric, achieve.org*)

3	a significant majority or all aspects of the criterion
2	some aspects of the criterion
1	none or nearly none of the aspects of the criterion

For shorter artifacts, Insight used the EQuIP Task Review Rubric for ELA, also used by ISBE. It assesses standard alignment, attention to text complexity and the shifts in the standards, and implementation support. Insight rated these artifacts with the following descriptors:

E	Most criteria checked
E/I	Many criteria checked but could use minor improvements
R	Some criteria checked
N	Task not recommended for instruction

ENGLISH

The following tables show artifact numbers (assigned by the order in which they were analyzed) and their 1-3 ratings on Criteria A-F on the English assessment criteria, or E-N ratings on the adapted EQuIP task rubric.

	Artifact number									
Criterion	26	40	41	42	43	44	59	60	61	62
A	3	3	3	3	3	3	3	1	3	1
B	1	1	1	2	2	2	2	2	2	1
C	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	3	3	3	3	3
E	3	3	3	3	3	2	2	2	2	2

Criterion	63	82	104	105
A	1			
B	1			
C	1			
D	3			
E	2	E/I	E/I	R

ENGLISH

The following short descriptions, by artifact number, provide additional information on the artifacts above.

26	AP Lang and Comp Common Interim A. No writing passages. One reading passage. Should increase cognitive complexity of added multiple choice question. Mostly Selected Response when it should only be 45%.
40	AP Lang and Comp Common Interim E. No writing passages. One reading passage. Should increase cognitive complexity of added multiple choice question. Mostly Selected Response when it should only be 45%.
41	AP Lang and Comp Common Interim D. No writing passages. One reading passage. Should increase cognitive complexity of added multiple choice question. Mostly Selected Response when it should only be 45%.
42	English I CIE. Lexile appropriate, texts worth reading. One fic one nonfic. 14 Selected Response. RI/RL .1 and .2. No linked passages. No writing.
43	English I CIC. Lexile appropriate, texts worth reading. One fic two nonfic. 14 Selected Response. RI/RL .1 and .2. No linked passages. No writing.
44	English II CIA. Has 30 pt writing prompt. Text worth reading. Long nonfic. 15 Selected Responses, variety of standards yet only 1 standard tagged in answer key.
59	English II CIE. Has 30 pt writing prompt. Text worth reading. Long nonfic. 15 Selected Responses, variety of standards. Might the interim include references to texts already read during unit of study (per standard 9)? Only standards Reading standards 1 and 3 assessed (or tagged)
60	IH Survey of Literature CIB. 5th grade lexile text, themes and content 6-8th grade. Only R.1 standard tagged. Has writing prompt and 12 Selected Responses. No references to other text, only 1 genre of text.
61	IH Survey of Literature CID. Lexile and themes appropriate to 9th grade. Only R.1 standard tagged to items. Has writing prompt and 12 Selected Responses. No references to other text, only 1 genre of text.
62	IH World Literature CIB. No text to read or react to. Only L.1 and RL.2, but there is no text. Has writing prompt and 12 Selected Responses. No references to any texts.
63	IH World Literature CID. No text to read or react to. Only L.1 and RL.2, but there is no text. Has writing prompt and 12 Selected Responses. No references to any texts.

82	English I formative. 810-1000L puts it at ~4th grade. WaPo OpEd = relevant themes. Task aligns with CCSS expectations for grade level and questions worth answering. Evidence-based writing, attention to key ideas and details, and structure of the argument. Builds disciplinary knowledge. E/I—most things checked, minor improvements may include bolstering text complexity with an excerpt from the Tech Times article.
104	English I, formative. Genocide Research project. The question or topic is one worth answering/exploring—genocide and its consequences. Needs explicit connection to anchor text, so not sure of grade-level appropriateness. The task is not specific about which texts to read/research, and expectations for scope, length, information source. No rubric attached to provide students with learning targets.
105	English I formative. Reading Check, Act 3, R & J. Stays within content of CCSS RL. Text is appropriate. The questions are recall and do not go deep into the themes, students not asked to evaluate or argue. Does not involve rich and rigorous speaking and/or writing. If using this quality text, enrich the prompts with argument and evidence.

SCIENCE

Insight used criteria derived from various EQuIP/NGSS science rubrics from Achieve and appropriate for the purposes of this Audit, to assess science assessments on a 3-point scale.

Five Science Assessment Criteria and 3-point scale

A. Assesses state science standards to provide evidence about students' achievement in science. Assessment requires students to use some understanding of Disciplinary Core Ideas to successfully complete it, and includes Reading and Writing for Science and Technical standards.
B. Assessment requires students to use at least one Science and Engineering Practice to successfully complete the task.
C. Assessment requires students to identify and interpret evidence and engage in scientific reasoning as they make sense of phenomena and address problems.
D. There are varied task types requiring a range of analytical thinking and cognitive complexity.
E. Majority of assessment cannot be answered without information from tasks or items, nor can the majority of the assessment's items be answered successfully by using rote knowledge.

3	a significant majority or all aspects of the criterion
2	some aspects of the criterion
1	none or nearly none of the aspects of the criterion

For labs or shorter artifacts, Insight adapted the EQuIP Task Review Rubrics, also used by ISBE. This rubric assesses standard alignment, attention cross-cutting concepts and science and engineering practices, and implementation support. Insight rated these artifacts with the following descriptors:

E	Most criteria checked
E/I	Many criteria checked but could use minor improvements
R	Some criteria checked
N	Task not recommended for instruction

SCIENCE

The following tables show artifact numbers (assigned by the order in which they were analyzed) and their 1-3 ratings on Criteria A-F on the science assessment criteria, or E-N ratings on the adapted EQuIP task rubric.

	Artifact number									
Criterion	20	22	23	24	27	28	29	30	31	32
A	3	2	2	3	2	2	2	2	2	2
B	3	3	2	3	3	1	1	3	3	1
C	2	2	2	3	1	1	1	1	1	1
D	3	3	3	3	1	2	2	2	1	2
E	2	1	2	3	1	1	1	2	2	2

Criterion	36	37	38	39	45	46	47	48	49	50
A	2	2	2	2	2	2	2	2	2	2
B	1	1	1	1	1	1	2	2	2	3
C	1	1	1	1	2	1	1	1	1	2
D	2	1	1	1	1	1	1	1	1	3
E	1	1	1	1	1	1	1	1	1	3

Criterion	51	52	73	74	75	76	77	85	88	89
A	2	2				2	2		2	2
B	1	1				3	3		3	3
C	1	1				3	3		2	2
D	1	1				3	3		3	1
E	1	1	E/I	E/I	E	3	3	R	3	3

Criterion	90	91	95	98	102	103
A	2	3	2	2		
B	3	3	1	3		
C	1	2	1	1		
D	1	2	1	2		
E	3	3	1	3	E	N

SCIENCE

The following short descriptions, by artifact number, provide additional information on the artifacts above.

20	Bio. Variety of item types and requires analysis of info provided by food webs. More than half is Selected Response, success for which would come from memorizing/content study vs. application. Open Response contains WHST and RI push to cite evidence.
22	Chemistry stoichiometry test. Driven by one lab and inclusive of many item types and SEPs. Gives step-by-step directions which may limit critical thinking and problem solving.
23	Bio cell test. Varied item types, includes Open Response. Mostly Selected Response and fill-in-the-blank. Open Response section 1 is strongest as it asks students "why" questions but before fill-in-the-blank section 2. Needs more writing but length is appropriate.
24	Chemistry density test. Item variety includes Open Response. Writing, graph reading, use of graphs and figures. Short Answers require problem solving using data presented.
27	Chemistry Unit II test. All Selected Response, no Open Response → no synthesis of dimensions or crosscutting, no reading or writing. Can succeed by studying/memorizing.
28	Chemistry ch I and II test. Open Response choice does not require reasoning and does not require citing from text evidence or use of sources. No use of SEPs.
29	Same artifact as #28
30	IH Chem unit 4. 15 Selected Response, 4 Short Answer and 6 matching. No Open Response. Most is from memorization/study. No significant reading/writing.
31	IH Chem Gas Laws Unit exam. 15 Selected Response. No Open Response.
32	Chem exam chapters 5-A. 27 SR. No Open Response → no synthesis of dimensions or crosscutting, no reading or writing. Some Selected Responses require distractor analysis. Use of graphics to answer questions.
36	Biology cell bio unit test A. Lengthy at 57 questions. 47 Selected Response, 6 matching, 4 Short Answer with some "why" prompting. No Open Response → no synthesis of dimensions or crosscutting, no reading or writing. No SEPs.

37	Bio. All Selected Response, no Open Response → no synthesis of dimensions or crosscutting, no reading or writing. Can succeed by studying/memorizing. Some use of graphs/charts/tables (nutrition label)
38	Bio, All Selected Response, no Open Response → no synthesis of dimensions or crosscutting, no reading or writing. Can succeed by studying/memorizing.
39	Bio, All Selected Response, no Open Response → no synthesis of dimensions or crosscutting, no reading or writing. Can succeed by studying/memorizing.
45	Biology Final. All Selected Response, no Open Response → no synthesis of dimensions or crosscutting, no reading or writing. Can succeed by studying/memorizing. Some critical analysis when referring to figures.
46	Biology final. All Selected Response, no Open Response → no synthesis of dimensions or crosscutting, no reading or writing. Can succeed by studying/memorizing.
47	IH Chem final. All Selected Response, no Open Response → no synthesis of dimensions or crosscutting, no reading or writing. Can succeed by studying/memorizing. Some math.
48	Chem final. All Selected Response, no Open Response → no synthesis of dimensions or crosscutting, no reading or writing. Can succeed by studying/memorizing. Some math.
49	Chem final. See #48.
50	Chem final. Strong variety. Many tandem items. No Open Response. Much of Selected Response is labeling. Some Selected Response is text-heavy (which is good).
51	Bio final. All Selected Response, no Open Response → no synthesis of dimensions or crosscutting, no reading or writing. Can succeed by studying/memorizing.
52	Bio final. All Selected Response, no Open Response → no synthesis of dimensions or crosscutting, no reading or writing. Can succeed by studying/memorizing.
73	IH Biology formative. Bellringer? To strengthen recall question #1, have students explain why the checks in the table should be where they are. This incorporates more SEPs and integrates the two questions.
74	IH Biology formative. Onion osmosis lab. Directions are simple and leave more room for exploration. Prompts encourage observation at first, "draw what you observe." Strengthen the culminating question by more directly connecting what is happening in observation to science terms being studied (e.g., osmosis) or asking students to explain why the cell structures behave the way they did using CER.

75	IH Biology formative. Microscope lab. Comprehensive. Instruction on use of hardware and why. Instruction on hardware connected to actual observations aligned with expectations of NGSS. Using SEPs and CCC between physical and life sciences. Post-lab questions ask students to argue using evidence.
76	IH Biology CIB. Lots of grade-level reading, problem-solving, and SEPs/CCCs. Selected Response and major Open Response/writing component. Taps into knowledge of ecosystems (Disciplinary Core Instruction). Needs to state which life science and other DCI standards this assesses.
77	IH Biology CID. Lots of grade-level reading, problem-solving, and SEPs/CCCs. SR and major OR/writing component. Taps into knowledge of ecosystems (Disciplinary Core Iinstruction). Needs to state which life science and other DCI standards this assesses.
85	Chemistry formative. Requires some understanding of chemistry DCI. No reading and writing standards used. No SEPs. No need to interpret evidence and engage in reasoning to solve problems—all ID'ing, defining, Selected Response. All short-answer task types. All can be answered without information provided by the task itself.
88	IH Chem, formative. Task is about understanding gas laws. Need more info about how their finding and then demonstrating an example of the gas law at work shows true understanding of, to quote the task, how "factors that affect the behavior of gases" interact. Biggest point total is explaining these factors: perhaps clarify that earning those points requires clear explanation of each variable and its expression in the example, why the formula divides, multiplies, adds, etc. Students perform several SEPs. Task pushes students to engage in scientific reasoning. Task types could vary (calculation, writing, explaining)—be clearer in the rubric what you want to see. Task results are dependent on information gleaned during the task.
89	IH Chem, formative. Students use understanding of moles and molar mass to complete. All calculations, no reading/writing. Yes SEPs (math). Less having to interpret evidence and reason, more calculation and use of formula (no having to answer why, or reasonableness, etc.). All calculation and showing work, without having to explain or make connections. Students need the info in the task to complete it.
90	Chem, formative. Students display understanding of molar mass formula. No reading/writing. Yes SEPs (math). No needing to interpret evidence, only applying calculation to formula. No varied task types. Need the information on the task to complete it.
91	Bio, formative. Lab. Students use understanding of macromolecule function. Students write to complete lab. Students analyze data, argue from evidence,

	and draw conclusions. They engage with evidence, questions are designed to lead them to reasoning. Questions are sequenced in a way that decreases student cognitive load. Most of the assessment can't be answered without info from the task.
95	IH Chem, formative. Students use understanding of nonpolar and polar covalent bonds, and their shapes. No writing. Little in the form of SEPs, reasoning, or interpreting evidence. All recall/short answer. Assessment can be answered without info from the task (recall).
98	Bio, formative. Tests understanding of scientific method, experimental design, graphing. No writing. Students use models and analyze/interpret data. Reasoning and sense-making less used than recall and identifying. Some variation in Selected Resopnse/match/order, no writing. Need the distractors to answer the questions
102	Chemistry, formative. As an ET, does it incorporate reading and writing, and an understanding of the Core Ideas? Yes. The SEP of supporting from evidence, yes. Interpreting evidence and engaging in scientific reasoning, yes. Appropriate analytical thinking and cognitive complexity for an ET. Needs the assessment info to complete it. Yes.
103	Bio, formative. Quiz. Some understanding of DCIs, no writing or reading necessary. No SEPs. No reasoning and not addressing a problem. All recall. All rote knowledge.

SOCIAL STUDIES

Insight used criteria derived from leading assessment evaluation tools, and appropriate for the purposes of this Audit, to assess social studies assessments on a 3-point scale. Evaluation tools of origin are linked.

Five Social Studies Assessment Criteria and 3-point scale

A. Texts, including primary sources, are worth reading. Texts reflect the quality of writing that is produced by authorities in the social sciences (*from Criteria for Procuring and Evaluating High-Quality Assessments, CCSSO*).
B. Informational texts and tasks associated with them reflect demands of shifts (citing evidence, building knowledge, text complexity) and standards (*from Criteria for Procuring and Evaluating High-Quality Assessments, CCSSO*)
C. Assessments and tasks should include purposefully linked passages or questions that allude to other accessible historical texts or sources (*from Assessment Evaluation Tool, achievethecore.org*)
D. Variety of item types to accurately assess a standard. Students should have the opportunity to write in response to high-quality texts and primary sources (*from Criteria for Procuring and Evaluating High-Quality Assessments, CCSSO*)
E. Items should be designed to elicit direct, observable evidence of the degree to which a student can independently demonstrate the key understandings relative to the demands of the social studies standard(s) (*from EQuIP rubric, achieve.org*)

3	a significant majority or all aspects of the criterion
2	some aspects of the criterion
1	none or nearly none of the aspects of the criterion

For shorter artifacts, Insight adapted the EQuIP Task Review Rubrics, also used by ISBE. This rubric assesses standard alignment, attention to teaching strategies and literacy strategies, and implementation support. Insight rated these artifacts with the following descriptors:

E	Most criteria checked
E/I	Many criteria checked but could use minor improvements
R	Some criteria checked
N	Task not recommended for instruction

SOCIAL STUDIES

The following tables show artifact numbers (assigned by the order in which they were analyzed) and their 1-3 ratings on Criteria A-F on the social studies assessment criteria, or E-N ratings on the adapted EQuIP task rubric.

	Artifact number									
Criterion	16	17	18	19	21	25	53	54	56	72
A	3	1	1	x	x	3	3	1	3	n/a
B	2	1	1	2	2	2	2	1	2	
C	x	x	x	x	x	x	x	x	x	
D	3	1	1	1	1	3	1	1	3	
E	2	1	1	3	3	2	1	1	2	

Criterion	83	86	87	111
A				1
B				1
C				x
D				2
E	R	N	N	2

SOCIAL STUDIES

The following short descriptions, by artifact number, provide additional information on the artifacts above.

16	USH Common Interim A. Lexile appropriate. Worth reading (nonfic). Only 1 reading passage, is there room for a short primary source?
17	USH interim from West. No texts. All Selected Response. All fact recall. No writing.
18	See #18.
19	IH USH unit test. No texts within the assessment. Mostly Selected Response (some matching). Selected Responses contain some graphs and some short primary source documents. Textbook is United States History from Prentice Hall 2013. Unclear lexile level.
21	IH USH unit test. See #19
25	USH common interim. See #16
53	IH USH final. Some map and graph reading. Read and respond to two texts, one lexile appropriate, the other 3-4th grade but primary source memoirs. No writing/essays.
54	USH final. 12 items, all Short Answer. No items gather evidence that a student can independently demonstrate key understanding of US History.
56	USH Common Interim F. Lexile appropriate. Worth reading (nonfic). Only 1 reading passage, is there room for a short primary source? Writing prompt "describe what happened" needs scope narrowed and greater specificity.
72	IH Global Studies formative. Graphic organizer to collect thoughts on 3 most important learnings, 2 questions needing answers, 1 connection with prior knowledge. No rating.
83	USH formative. Questions ask students to write arguments for or against. Good why questions. Text is below grade level and packaged by teacher, rather than primary source document.
86	USH bellringer (formative). Task aligns with History standards but doesn't require analysis of text to answer. Recall questions. No challenging reading or writing. As bellringer—activation of current lesson nor connection to past lessons not apparent.
87	USH bellringer (formative). Task aligns with History standards but doesn't require analysis of text to answer. Recall questions. No challenging reading

	or writing. As bellringer—activation of current lesson nor connection to past lessons not apparent.
111	IH Global Studies, semester one final. 60 points. Half are Selected Response and facts/recall. Other half is choice of one of three essays. No texts in the assessment. Assessment does not reflect shifts in the standards except for directions to cite evidence (no text from which to cite it). Students can allude to other texts, but no opportunity to write in response to explicitly stated high-quality texts. Selected Response items elicit direct understanding relative to demands of the standard, but without a text, the students respond to the Open Responses based on memorized knowledge only.

HEALTH & WELLNESS

With the exception of two IH Health/Wellness common interims, Insight adapted the EQuIP Task Review Rubrics, also used by ISBE, to assess the remaining nine Health/Wellness artifacts, which were all from Driver's Education or IH Driver's Education. This rubric assesses standard alignment, attention to teaching strategies and literacy strategies, and implementation support. Insight rated these artifacts with the following descriptors:

E	Most criteria checked
E/I	Many criteria checked but could use minor improvements
R	Some criteria checked
N	Task not recommended for instruction

Two IH Health/Wellness assessments were rated on the five science Assessment Criteria:

A. Assesses state science standards to provide evidence about students' achievement in science. Assessment requires students to use some understanding of Disciplinary Core Ideas to successfully complete it, and includes Reading and Writing for Science and Technical standards.
B. Assessment requires students to use at least one Science and Engineering Practice to successfully complete the task.
C. Assessment requires students to identify and interpret evidence and engage in scientific reasoning as they make sense of phenomena and address problems.
D. There are varied task types requiring a range of analytical thinking and cognitive complexity.
E. Majority of assessment cannot be answered without information from tasks or items, nor can the majority of the assessment's items be answered successfully by using rote knowledge.

HEALTH & WELLNESS

The following tables show artifact numbers (assigned by the order in which they were analyzed) and their 1-3 ratings on Criteria A-F on the social science criteria, or E-N ratings on the adapted EQuIP task rubric.

	Artifact number										
Criterion	9	10	11	12	55	57	80	81	100	101	109
A	E/I	E/I	E/I	E/I	E/I	E/I	3	3	E/I	E/I	E/I
B							3	3			
C							3	3			
D							3	3			
E							3	3			

HEALTH & WELLNESS

The following short descriptions, by artifact number, provide additional information on the artifacts above.

9	IH DE Test. All from Rules of the Road.
10	IH DE formative assessments. From required chapters of Rules of the Road.
11	DE formative assessments
12	DE formative assessments. See #11.
55	DE Final, west. All from Rules of the Road.
57	IH DE final. All from Rules of the Road.
80 (using science criteria)	CIA IH Health/Wellness II. Multiple choice and populate a table. Read graphs and use mathematics to solve problems. 2 IL PE standards. (19-24 are covered in this class, according to S&S). CIA, CIB, and CIC are the same.
81 (using science criteria)	CIC IH Health/Wellness II. See #80. CIA, CIB, and CIC are the same.
100	DE formative. Lots of short answer, content and answer format matches the permit test and content of Rules of the Road?
101	DE formative. Lots of short answer, content and answer format matches the permit test and content of Rules of the Road?
109	DE. Study guide. Lots of short answer, content and answer format matches the permit test and content of Rules of the Road?

WORLD LANGUAGES

For all eleven World Languages (Spanish I or IH Spanish I) artifacts, Insight adapted the EQuIP Task Review Rubric, also used by ISBE. It assesses standard alignment, attention to reading and writing, and attention to speaking and listening. Insight rated these artifacts with the following descriptors:

E	Most criteria checked
E/I	Many criteria checked but could use minor improvements
R	Some criteria checked
N	Task not recommended for instruction

The following table shows artifact numbers (assigned by the order in which they were analyzed) and their E-N ratings on the adapted EQuIP task rubric.

	Artifact number										
	13	14	15	33	34	35	58	78	79	106	108
Rating	E	E	E/I	E/I	E/I	E	E	E	E	E	E

WORLD LANGUAGES

The following short descriptions, by artifact number, provide additional information on the artifacts above.

13	Spanish I common interim E. ACTFL standards on interpersonal and cultural communication. English and Writing standards on citing evidence from text. Two texts, one dialogue and other written comm.
14	IH Spanish I project (formative). Dream House project. Not sure if text/video is grade level. Assuming it is, this formative contains balance of modalities, variety of ways to assess (speaking and creating) and allows for direct observable evidence of key understandings relative to speaking/listening and writing standards.
15	Spanish I Final conference. Clear and coherent writing standards + speaking and listening standards. At the end of Saludos Unit (1). Contains all but practicing of the verb ser.
33	Spanish I test. Has text (dialog with greetings) worth reading. All matching, fill in the blank, and SR. No OR/writing. Mixes real world content with fictional dialog. Uses figures and pictures.
34	See #33.
35	IH Spanish I project (formative). Mi familia project. this formative contains balance of modalities, variety of ways to assess (speaking and creating) and allows for direct observable evidence of key understandings relative to speaking/listening and writing standards.
58	Spanish I final. Item variety, different genres of text.
78	IH Spanish I CIB. See #13
79	IH Spanish I CID. See #13
106	Spanish I formative. Verbs quiz. Task aligns with standards, subject-verb alignment is within the context of a story/paragraph. Story/paragraph provides opportunities to build knowledge while practicing. Variety of task types.
108	Spanish I formative. Question words quiz. Task aligns with standards, question words choices within the context of a story/paragraph. Story/paragraph provides opportunities to build knowledge while practicing. Variety of task types, including multiple-select Selected Response.

UNIT AND TASK ANALYSIS DOCUMENTS

Grade: **Literacy Lesson/Unit Title:** **Overall Rating:**

I. Alignment to the Depth of the CCSS	II. Key Shifts in the CCSS	III. Instructional Supports	IV. Assessment
The lesson/unit aligns with the letter and spirit of the CCSS: ○ Targets a set of grade-level CCSS ELA/Literacy standards. ○ Includes a clear and explicit purpose for instruction. ○ Selects text(s) that measure within the grade-level text complexity band and are of sufficient quality and scope for the stated purpose (e.g., presents vocabulary, syntax, text structures, levels of meaning/purpose, and other qualitative characteristics similar to CCSS grade-level exemplars in Appendices A & B). *A unit or longer lesson should:* ○ Integrate reading, writing, speaking and listening so that students apply and synthesize advancing literacy skills. ○ (Grades 3-5) Build students' content knowledge and their understanding of reading and writing in social studies, the arts, science or technical subjects through the coherent selection of texts.	*The lesson/unit addresses key shifts in the CCSS:* ○ **Reading Text Closely:** Makes reading text(s) closely, examining textual evidence, and discerning deep meaning a central focus of instruction. ○ **Text-Based Evidence:** Facilitates rich and rigorous evidence-based discussions and writing about common texts through a sequence of specific, thought-provoking, and text-dependent questions (including, when applicable, questions about illustrations, charts, diagrams, audio/video, and media). ○ **Writing from Sources:** Routinely expects that students draw evidence from texts to produce clear and coherent writing that informs, explains, or makes an argument in various written forms (e.g., notes, summaries, short responses, or formal essays). ○ **Academic Vocabulary:** Focuses on building students' academic vocabulary in context throughout instruction. *A unit or longer lesson should:* ○ **Increasing Text Complexity:** Focus students on reading a progression of complex texts drawn from the grade-level band. Provide text-centered learning that is sequenced, scaffolded and supported to advance students toward independent reading of complex texts at the CCR level. ○ **Building Disciplinary Knowledge:** Provide opportunities for students to build knowledge about a topic or subject through analysis of a coherent selection of strategically sequenced, discipline-specific texts. ○ **Balance of Texts:** Within a collection of grade-level units a balance of informational and literary texts is included according to guidelines in the CCSS (p. 5). ○ **Balance of Writing:** Include a balance of on-demand and process writing (e.g., multiple drafts and revisions over time) and short, focused research projects, incorporating digital texts where appropriate.	*The lesson/unit is responsive to varied student learning needs:* ○ Cultivates student interest and engagement in reading, writing and speaking about texts. ○ Addresses instructional expectations and is easy to understand and use. ○ Provides *all* students with multiple opportunities to engage with text of appropriate complexity for the grade level; includes appropriate scaffolding so that students directly experience the complexity of the text. ○ Focuses on challenging sections of text(s) and engages students in a productive struggle through discussion questions and other supports that build toward independence. ○ Integrates appropriate supports in reading, writing, listening and speaking for students who are ELL, have disabilities, or read well below the grade level text band. ○ Provides extensions and/or more advanced text for students who read well above the grade level text band. *A unit or longer lesson should:* ○ Include a progression of learning where concepts and skills advance and deepen over time (*may be more applicable across the year or several units*). ○ Gradually remove supports, requiring students to demonstrate their independent capacities (*may be more applicable across the year or several units*). ○ Provide for authentic learning, application of literacy skills, student-directed inquiry, analysis, evaluation and/or reflection. ○ Integrate targeted instruction in such areas as grammar and conventions, writing strategies, discussion rules and all aspects of foundational reading for grades 3-5. ○ Indicate how students are accountable for independent reading based on student choice and interest to build stamina, confidence and motivation (*may be more applicable across the year or several units*). ○ Use technology and media to deepen learning and draw attention to evidence and texts as appropriate.	*The lesson/unit regularly assesses whether students are mastering standards-based content and skills:* ○ Elicits direct, observable evidence of the degree to which a student can independently demonstrate the major targeted grade-level CCSS standards with appropriately complex text(s). ○ Assesses student proficiency using methods that are unbiased and accessible to all students. ○ Includes aligned rubrics or assessment guidelines that provide sufficient guidance for interpreting student performance. *A unit or longer lesson should:* ○ Use varied modes of assessment, including a range of pre-, formative, summative and self-assessment measures.
Rating: 3 2 1 0	**Rating: 3 2 1 0**	**Rating: 3 2 1 0**	**Rating: 3 2 1 0**

EQuIP Rubric for Lessons & Units: ELA/Literacy (Grades 3-5) and ELA (Grades 6-12)

Directions: The Quality Review Rubric provides criteria to determine the quality and alignment of lessons and units to the Common Core State Standards (CCSS) in order to: (1) Identify exemplars/ models for teachers' use within and across states; (2) provide constructive criteria-based feedback to developers; and (3) review existing instructional materials to determine what revisions are needed.

Step 1 – Review Materials

- Record the grade and title of the lesson/unit on the recording form.
- Scan to see what the lesson/unit contains and how it is organized.
- Read key materials related to instruction, assessment and teacher guidance.
- Study and measure the text(s) that serves as the centerpiece for the lesson/unit, analyzing text complexity, quality, scope, and relationship to instruction.

Step 2 – Apply Criteria in Dimension I: Alignment

- Identify the grade-level CCSS that the lesson/unit targets.
- Closely examine the materials through the "lens" of each criterion.
- Individually check each criterion for which clear and substantial evidence is found.
- Identify and record input on specific improvements that might be made to meet criteria or strengthen alignment.
- Enter your rating 0 – 3 for Dimension I: Alignment

Note: Dimension I is non-negotiable. In order for the review to continue, a rating of 2 or 3 is required. If the review is discontinued, consider general feedback that might be given to developers/teachers regarding next steps.

Step 3 – Apply Criteria in Dimensions II – IV

- Closely examine the lesson/unit through the "lens" of each criterion.
- Record comments on criteria met, improvements needed and then rate 0 – 3.

When working in a group, individuals may choose to compare ratings after each dimension or delay conversation until each person has rated and recorded their input for the remaining Dimensions II – IV.

Step 4 – Apply an Overall Rating and Provide Summary Comments

- Review ratings for Dimensions I – IV adding/clarifying comments as needed.
- Write summary comments for your overall rating on your recording sheet.
- Total dimension ratings and record overall rating E, E/I, R, N – adjust as necessary.

If working in a group, individuals should record their overall rating prior to conversation.

Step 5 – Compare Overall Ratings and Determine Next Steps

- Note the evidence cited to arrive at final ratings, summary comments and similarities and differences among raters. Recommend next steps for the lesson/unit and provide recommendations for improvement and/or ratings to developers/teachers.

Additional Guidance for ELA/Literacy – When selecting text(s) that measure within the grade-level text complexity band and are of sufficient quality and scope for the stated purpose, see *The Common Core State Standards in English Language Arts/Literacy* at www.corestandards.org/ELA-Literacy; and the *Supplement for Appendix A: New Research on Text Complexity as well as Quantitative and Qualitative Measures* at www.achievethecore.org/steal-these-tools/text-complexity. See *The Publishers' Criteria* for Grades K-2 and the same for Grades 3-12 at www.achievethecore.org/steal-these-tools.

Rating Scales

Note: Rating for Dimension I: Alignment is non-negotiable and requires a rating of 2 or 3. If rating is 0 or 1 then the review does not continue.

Rating Scale for Dimensions I, II, III, IV:	**Overall Rating for the Lesson/Unit:**
3: Meets most to all of the criteria in the dimension	**E: Exemplar –** Aligned and meets most to all of the criteria in dimensions II, III, IV **(total 11 – 12)**
2: Meets many of the criteria in the dimension	**E/I: Exemplar *if* Improved –** Aligned and needs some improvement in one or more dimensions **(total 8 – 10)**
1: Meets some of the criteria in the dimension	**R: Revision Needed** – Aligned partially and needs significant revision in one or more dimensions **(total 3 – 7)**
0: Does not meet the criteria in the dimension	**N: Not Ready to Review –** Not aligned and does not meet criteria **(total 0 – 2)**
Descriptors for Dimensions I, II, III, IV:	**Descriptors for Overall Rating:**
3: Exemplifies CCSS Quality – meets the standard described by criteria in the dimension, as explained in criterion-based observations.	**E: Exemplifies CCSS Quality –** Aligned and exemplifies the quality standard and exemplifies most of the criteria across Dimensions II, III, IV of the rubric.
2: Approaching CCSS Quality – meets many criteria but will benefit from revision in others, as suggested in criterion-based observations.	**E/I: Approaching CCSS Quality –** Aligned and exemplifies the quality standard in some dimensions but will benefit from some revision in others.
1: Developing toward CCSS Quality – needs significant revision, as suggested in criterion-based observations.	**R: Developing toward CCSS Quality –** Aligned partially and approaches the quality standard in some dimensions and needs significant revision in others.
0: Not representing CCSS Quality – does not address the criteria in the dimension.	**N: Not representing CCSS Quality –** Not aligned and does not address criteria.

EQuIP Task Review Rubric – ELA

Task Title_____ **Grade** _____ **Date**_____ **Rating:**

Targeted Standards_____

I. Alignment to the CCSS	II. Attention to Text Complexity & the Instructional Shifts
The task clearly aligns with one or more CCSS. ☐ The task directly aligns with the content and expectations of one or more CCSS. ☐ The task addresses a question worth answering and requires analysis of a grade-level text and/or topic. ☐ Selects text(s) that measure within the grade-level text complexity band.	*The task supports the key shifts that are reflected in the CCSS.* ☐ The task involves reading grade-level texts closely, attending to key ideas and details, language, and/or craft and structure. ☐ The task involves rich and rigorous evidence-based speaking and/or writing. ☐ The task provides opportunities to build important disciplinary knowledge and vocabulary.
Notes and observations regarding alignment and attention to the instructional shifts of the CCSS:	

III. Support for Implementation
The task includes relevant supporting information or materials that ensure effective administration of the task and evaluation of student thinking. ☐ The task can be used to elicit direct, observable evidence of the degree to which each student can demonstrate the skills and knowledge addressed in the targeted CCSS. ☐ Supporting materials include answer keys, rubrics, and/or scoring guidelines that are clearly connected to the targeted CCSS and provide sufficient guidance for interpreting student performance. ☐ The task's prompts and directions provide sufficient guidance for the teacher to administer it effectively and for the students to complete it successfully. ☐ The task is accessible to and appropriate for <u>all</u> learners, including students who are English language learners or are working below or above grade level. ☐ The task cultivates student interest and/or engagement in reading, writing, and speaking about text.
Notes and observations regarding support features that may be required for effective administration of the task:

Rating Descriptors:

E: Most criteria are checked, <u>including all three in Dimension I</u> and those that are appropriate for the task's purpose in Dimension II and III. The task is likely to promote successful learning and/or assessment of the skills and knowledge required in the targeted CCSS.

E/I: Many criteria are checked. The task is aligned to the CCSS and has potential but could benefit from some minor improvements.

R: Some criteria are checked. The task has potential but needs significant revision to be considered effective.

N: The task is not recommended for instruction and/or assessment of the CCSS.

Assessment Evaluation Tool (AET)

ELA/Literacy, Grades 3–12

Assessment Evaluation Tool

ELA/Literacy, Grades 3–12

This ELA/literacy AET is designed to help educators determine whether or not assessments and sets of assessments are aligned to the Shifts and major features of the Common Core State Standards (CCSS). The substantial instructional Shifts (http://www.corestandards.org/other-resources/key-shifts-in-english-language-arts/) at the heart of the Common Core State Standards are:

- **Complexity**: Regular practice with complex text and its academic language
- **Evidence**: Reading, writing, and speaking grounded in evidence from text, both literary and informational
- **Knowledge**: Building knowledge through content-rich non-fiction

The AET draws directly from the following documents:

- Common Core State Standards for English Language Arts & Literacy in History/Social Studies, Science, and Technical Subjects (http://www.corestandards.org/ELA-Literacy/)
- Publishers' Criteria for the Common Core State Standards in ELA/Literacy grades 3 – 12 (http://corestandards.org/assets/Publishers_Criteria_for_3-12.pdf)
- Supplement to Appendix A of the Common Core State Standards for ELA/Literacy: New Research on Text Complexity (www.corestandards.org/assets/E0813_Appendix_A_New_Research_on_Text_Complexity.pdf)

When to use the AET

1. Purchasing assessments: Many factors go into local purchasing decisions. Alignment to the Standards is a critical factor to consider. The AET is designed to evaluate alignment of assessments and sets of assessments to the Shifts and the major features of the CCSS. It also provides suggestions of additional indicators to consider in the assessment evaluation and purchasing process.
2. Evaluating assessments in use: The AET can be used to analyze the degree of alignment of existing assessments and sets of assessments and help to highlight specific, concrete flaws in alignment. Even where assessments currently in use fail to meet one or more of these criteria, the pattern of failure is likely to be informative. States and districts can use the evaluation to create a thoughtful plan to modify assessments and sets of assessments in such a way that they better meet the requirements of the Standards.
3. Developing assessments: This tool can be used to provide guidance for and evaluation of alignment for creating locally developed assessments and sets of assessments. Those developing new aligned assessments should use the criteria within the AET to guide test blueprint construction, item specifications development, and item evaluation procedures.

Who Uses the AET

The AET is designed for use by educators and administrators including content specialists, assessment specialists, administrators and educators at the school, district or state level. Evaluating assessments and sets of assessments requires both subject-matter and technical expertise. Evaluators should be well versed in the Standards (http://www.corestandards.org/ELA-Literacy/) for all grades in which assessments are being evaluated. Evaluators also should be familiar with the substantial instructional Shifts (http://www.corestandards.org/other-resources/key-shifts-in-english-language-arts/) of Complexity, Evidence and Knowledge that are listed above. If possible, it would be helpful if at least one member of the evaluation team is well versed in ELA/literacy assessment.

Getting Started

Prior to Evaluation

Assemble all of the materials necessary for the evaluation, e.g., test forms, test blueprints, test item metadata, item bank summaries, sample score reports. It is essential to have materials for all grades covered by the assessment program, as some criteria cannot be rated without having access to each grade. In addition, each evaluator should have a reference copy of the Common Core State Standards for ELA/Literacy and the Publishers' Criteria for the Common Core State Standards in ELA/Literacy grades 3 – 12.

Sections 1 – 3 below should be completed to produce a comprehensive picture of the strengths and weaknesses of the assessments under evaluation. Information about areas in need of improvement should be shared with internal and external stakeholders.

Navigating the Tool

The AET contains criteria for five ELA/literacy domains: Reading, Writing, Language, and Speaking and Listening. Assessments do not have to contain all of the ELA/literacy domains in order to be evaluated with the AET or to align with the CCSS. Choose the Non-Negotiables and/or Alignment Criteria that apply to the assessments being evaluated.

If reading is being assessed*, begin with Section 1: Non-Negotiable Alignment Criteria (p. 4).

- The Non-Negotiable Alignment Criteria must each be met in full for reading assessments to be considered aligned to the Shifts and the major features of the Common Core State Standards. Each Non-Negotiable Alignment Criterion has three metrics associated with it; every one of these metrics must be met in order for the criterion as a whole to be met.
- Examine the relevant materials and use evidence to rate the materials against each criterion and its associated metrics.
- Record and explain the evidence upon which the rating is based.

Continue to Section 2: Alignment Criteria (p. 14).

- The Alignment Criteria for the domains covered by the assessment program under evaluation must each be met for materials to be considered aligned to the Shifts and the major features of the Common Core State Standards. Each Alignment Criterion has two or more metrics associated with it; a specific number of these metrics must be met or partially met in order for the criterion as a whole to be met.
- The domains covered within the Alignment Criteria section are: Reading, Writing, Language, and/or Speaking and Listening.
- Examine the materials in relation to the relevant criteria, assigning each metric a point value. Rate each criterion as "Meets" or "Does Not Meet" based on the number of points assigned. The more points the materials receive on the alignment criteria, the better they are aligned.
- Record and explain the evidence upon which the rating is based.

Complete Section 3: Evaluation Summary (p. 43).

- Compile all of the results from Sections 1 and 2 to determine if the assessments are aligned to the Shifts and major features of the CCSS.

Proceed to Section 4: Indicators of Quality (p. 45).

- Indicators of Quality are important considerations that will help evaluators better understand the overall quality of an assessment program. These considerations are not criteria for alignment to the CCSS, but they provide valuable information about additional program characteristics, such as ensuring accessibility for all students. Evaluators may want to add their own indicators to the examples provided.

NOTE: The word "text" has been used to apply to written, audio, video, and quantitative stimuli. The AET should be applied to non-print materials as appropriate.

* It is assumed that reading will be a significant component of most assessment systems subject to evaluation. When an assessment does not include Reading, the Alignment Criteria for the domains being evaluated (Writing, Language, Speaking and Listening) should be used.

Directions for Non-Negotiable 1

Reading – Complexity and Quality of Texts

Non-Negotiable 1: Texts are worthy of student time and attention; they have the appropriate level of complexity for the grade, according to both quantitative and qualitative analyses of text complexity.

Required Materials

- The texts in the test forms for each grade level or (for an item bank) a random sample of texts for each grade level
- Metadata accompanying the texts, especially quantitative and qualitative analyses of text complexity and copyright acknowledgements

Rating this Criterion

The assessments should be rated for each of the following three metrics as Meets or Does Not Meet. If any one of the metrics is rated as Does Not Meet, then the assessments fail Non-Negotiable 1. If all metrics are rated as Meets, the assessments pass this Non-Negotiable.

Whether the assessments are rated as Meets or Does Not Meet, provide specific examples of evidence in support of the ratings, including evidence of any specific gaps in the assessments.

Non-Negotiable 1

Reading – Complexity and Quality of Texts

Metric

NN Metric 1A:
At least 90% of texts used for assessment are placed within the grade band indicated by a quantitative analysis, with the average complexity of texts increasing grade-by-grade. Exceptions—in which the text is placed above the indicated grade band—are usually reserved for literary texts in the upper grades. When materials are published, the quantitative data accompany the materials.

Procedure for Evaluation

Every text should be accompanied by data from at least one research-based quantitative tool for grade band placement (poetry and drama excepted). The same tool(s) should be used consistently across the grade levels.

If quantitative data is not available, evaluators should obtain a Lexile or other rating for the text (see http://achievethecore.org/text-complexity).

For each grade, examine the metadata or other explanatory materials accompanying either the texts on the test form(s) or a representative sample of at least three literary and three informational texts from the item bank.

Make a list of each text title and the grade to which it has been assigned; group by grade band. Note the grade band indicated by the quantitative tool(s) and the actual grade band placement.

Calculate an overall percentage of the texts that have been placed at or below the grade band indicated by the quantitative data, allowing exceptions for literary texts as appropriate.

Evidence

Rating

☐ **Meets**

☐ **Does Not Meet**

Non-Negotiable 1

Reading – Complexity and Quality of Texts

Metric

NN Metric 1B:
At least 90% of texts used for assessment are placed within the grade level indicated by a qualitative analysis. When materials are published, the qualitative analysis accompanies the materials.

Procedure for Evaluation

Every text should be accompanied by a qualitative analysis for grade level placement (including poetry and drama).

If a qualitative analysis is not available, evaluators should do a brief analysis using a format like the one at http://achievethecore.org/qualitative-text-analysis.

For each grade, examine the qualitative analyses in the metadata or other explanatory materials accompanying the same texts from Non-Negotiable 1A above. Note the grade level indicated by the qualitative tools and the actual grade level placement.

Calculate an overall percentage of the texts that have been placed at the grade level indicated by the qualitative analysis.

Evidence

Rating

☐ **Meets**

☐ **Does Not Meet**

Non-Negotiable 1

Reading – Complexity and Quality of Texts

Metric

NN Metric 1C:
At least 95% of texts used for assessment are of publishable quality—preferably previously published but at minimum edited by professional publication editors (not only assessment editors). History/social studies and science/technical texts, specifically, reflect the quality of writing that is produced by authorities in the particular academic discipline.

Procedure for Evaluation

All texts should be high quality and content rich—worthy of student attention. Nearly all texts should be previously published rather than "commissioned" because published texts have been selected and edited by professional publication editors.

For each grade, examine the metadata or other explanatory materials accompanying the same texts from Non-Negotiable 1A above.

Look for an acknowledgment line for each text (usually found at the front of the test booklet or below the text), which cites an author or publisher and date of publication, or look for a statement that the text has been edited by a professional publication editor.

Label the texts that are accompanied by an acknowledgment line or are shown to have been edited professionally.

Identify any texts that do not represent quality literary or informational writing.

Calculate the percentage of texts that are not of publishable quality.

Evidence

Rating

☐ **Meets**

☐ **Does Not Meet**

Non-Negotiable 1

Reading – Complexity and Quality of Texts

Non-Negotiable 1: Texts are worthy of student time and attention; they have the appropriate level of complexity for the grade, according to both quantitative and qualitative analyses of text complexity.

Rating for Non-Negotiable 1

If all three metrics above were rated as Meets, then rate Non-Negotiable 1 as Meets. If one or more of the metrics were rated as Does Not Meet, then rate Non-Negotiable 1 as Does Not Meet. Check the final rating.

Then, briefly describe the strengths and weaknesses of these materials in light of this Criterion.

Rating

☐ **Meets**

☐ **Does Not Meet**

Strengths / Weaknesses:

Before moving to Non-Negotiable 2, record the final Meets or Does Not Meet rating in the Evaluation Summary on Page 43.

Directions for Non-Negotiable 2

Reading – Text-Dependent and Standards-Based Questions

Non-Negotiable 2: High-quality reading test questions are text-dependent and Standards-based; they require students to read closely, find the answers within the text, and use textual evidence to support responses.

Required Materials

- The test questions in the test forms for each grade level or (for an item bank) a representative sample of test questions
- Metadata accompanying the test questions, showing the alignment of each question to the CCSS

Rating this Criterion

The assessments should be rated for each of the following three metrics as Meets or Does Not Meet. If any one of the metrics is rated as Does Not Meet, then the assessments fail Non-Negotiable 2. If all metrics are rated as Meets, the assessments pass this Non-Negotiable.

Whether the assessments are rated as Meets or Does Not Meet, provide specific examples of evidence in support of the ratings, including evidence of any specific gaps in the assessments.

Non-Negotiable 2

Reading – Text-Dependent and Standards-Based Questions

Metric

NN Metric 2A:
At least 90% of the questions are text dependent: they require close reading and analysis of the text, focus on its central ideas and important particulars, and require answers based on what is in (not outside) the text.

Procedure for Evaluation

Questions should require thoughtful reading of the text, not just skimming or superficial consideration. As a set, questions should enable students to demonstrate deep understanding of the unique aspects of the text. Students should be able to answer the questions correctly without prior knowledge. Questions should be derived from a reading text (i.e., not "stand alone" questions).

For each grade, examine either the test questions on the test form(s) or a representative sample of at least 15 questions based on literary texts and 15 based on informational texts per grade in the item bank.

Identify the questions that do not meet this metric: List the sequence numbers of any questions that do not require close reading and analysis, e.g., the questions assess simple recall or minor textual points. List the sequence numbers of any questions that, as a set, focus on peripheral aspects of the text, failing to permit students to demonstrate deep understanding of the text. List the sequence numbers of any questions that call on students' prior knowledge or are "stand-alone" questions.

Calculate percentages of test questions that do not meet the metric.

Evidence

Rating

☐ **Meets**

☐ **Does Not Meet**

Non-Negotiable 2

Reading – Text-Dependent and Standards-Based Questions

Metric

NN Metric 2B:
At least 90% of test questions reflect the range of cognitive demand required by the Standards.

Procedure for Evaluation

At every grade level, the Standards should be assessed with items that reflect a range of rigor and cognitive demand, depending on the requirements of individual Standards. Questions should reflect this range at each grade, always avoiding simple recall or surface analysis.

For each grade, examine the same test questions from Non-Negotiable 2A above.

List the sequence numbers of any questions that do not rise to the range of cognitive demand or rigor required by individual Standards.

Calculate a percentage of test questions that do not meet this metric.

Evidence

Rating

☐ **Meets**

☐ **Does Not Meet**

Non-Negotiable 2

Reading – Text-Dependent and Standards-Based Questions

Metric

NN Metric 2C:
At least 90% of test questions assess the specifics of the Standards at each grade level (not just the Anchor Standards) and do not employ "generic" answer choices applicable to any text.

Procedure for Evaluation

Questions should assess the specific requirements delineated by the Standards. For example, if a Standard requires a focus on two central ideas, two ideas should be assessed; if a Standard calls for the meaning of figurative language, meaning should be assessed, not literary terms like *metaphor* or *personification*.

Questions should not be aligned only to Anchor Standards. Multiple-choice or technology-enhanced items should be text-specific, not relying on "generic" choices (e.g., "to inform," "to persuade," "to entertain") that could be used for any text. Not every Standard must be assessed with every text.

For each grade, examine the test questions assembled under Non-Negotiable 2A above, along with their metadata. Identify the questions that do not meet this metric: List the sequence numbers of any questions that fail to assess the specific requirements of the Standards at the grade level. List the sequence numbers of any questions that are aligned only to the Anchor Standards. List the sequence numbers of any questions that provide "generic" answer choices that could be used for any text.

Calculate percentages of questions that do not meet the metric.

Evidence

Rating

☐ **Meets**

☐ **Does Not Meet**

Non-Negotiable 2

Reading – Text-Dependent and Standards-Based Questions

Non-Negotiable 2: High-quality reading test questions are text-dependent and Standards-based; they require students to read closely, find the answers within the text, and use textual evidence to support responses.

Rating for Non-Negotiable 2

If all three metrics above were rated as Meets, then rate Non-Negotiable 2 as Meets. If one or more of the metrics were rated as Does Not Meet, then rate Non-Negotiable 2 as Does Not Meet. Check the final rating.

Then, briefly describe the strengths and weaknesses of these materials in light of this Criterion.

Rating

☐ **Meets**

☐ **Does Not Meet**

Strengths / Weaknesses:

Before moving to the Alignment Criteria, record the final Meets or Does Not Meet rating in the Evaluation Summary on Page 43.

Now continue by evaluating the Alignment Criteria 1-4 for Reading.

Grade: **Mathematics Lesson/Unit Title:** ***EQuIP Rubric for Lessons & Units: Mathematics*** **Overall Rating:**

I. Alignment to the Depth of the CCSS	II. Key Shifts in the CCSS	III. Instructional Supports	IV. Assessment
The lesson/unit aligns with the letter and spirit of the CCSS: ○ Targets a set of grade-level CCSS mathematics standard(s) to the full depth of the standards for teaching and learning. ○ Standards for Mathematical Practice that are central to the lesson are identified, handled in a grade-appropriate way, and well connected to the content being addressed. ○ Presents a balance of mathematical procedures and deeper conceptual understanding inherent in the CCSS.	*The lesson/unit reflects evidence of key shifts that are reflected in the CCSS:* ○ **Focus:** Lessons and units targeting the major work of the grade provide an especially in-depth treatment, with especially high expectations. Lessons and units targeting supporting work of the grade have visible connection to the major work of the grade and are sufficiently brief. Lessons and units do not hold students responsible for material from later grades. ○ **Coherence:** The content develops through reasoning about the new concepts on the basis of previous understandings. Where appropriate, provides opportunities for students to connect knowledge and skills within or across clusters, domains and learning progressions. ○ **Rigor:** Requires students to engage with and demonstrate challenging mathematics with appropriate balance among the following: – **Application:** Provides opportunities for students to independently apply mathematical concepts in real-world situations and solve challenging problems with persistence, choosing and applying an appropriate model or strategy to new situations. – **Conceptual Understanding:** Develops students' conceptual understanding through tasks, brief problems, questions, multiple representations and opportunities for students to write and speak about their understanding. – **Procedural Skill and Fluency:** Expects, supports and provides guidelines for procedural skill and fluency with core calculations and mathematical procedures (when called for in the standards for the grade) to be performed quickly and accurately.	*The lesson/unit is responsive to varied student learning needs:* ○ Includes clear and sufficient guidance to support teaching and learning of the targeted standards, including, when appropriate, the use of technology and media. ○ Uses and encourages precise and accurate mathematics, academic language, terminology and concrete or abstract representations (e.g., pictures, symbols, expressions, equations, graphics, models) in the discipline. ○ Engages students in productive struggle through relevant, thought-provoking questions, problems and tasks that stimulate interest and elicit mathematical thinking. ○ Addresses instructional expectations and is easy to understand and use. ○ Provides appropriate level and type of scaffolding, differentiation, intervention and support for a broad range of learners. – Supports diverse cultural and linguistic backgrounds, interests and styles. – Provides extra supports for students working below grade level. – Provides extensions for students with high interest or working above grade level. *A unit or longer lesson should:* ○ Recommend and facilitate a mix of instructional approaches for a variety of learners such as using multiple representations (e.g., including models, using a range of questions, checking for understanding, flexible grouping, pair-share). ○ Gradually remove supports, requiring students to demonstrate their mathematical understanding independently. ○ Demonstrate an effective sequence and a progression of learning where the concepts or skills advance and deepen over time. ○ Expect, support and provide guidelines for procedural skill and fluency with core calculations and mathematical procedures (when called for in the standards for the grade) to be performed quickly and accurately.	*The lesson/unit regularly assesses whether students are mastering standards-based content and skills:* ○ Is designed to elicit direct, observable evidence of the degree to which a student can independently demonstrate the targeted CCSS. ○ Assesses student proficiency using methods that are accessible and unbiased, including the use of grade-level language in student prompts. ○ Includes aligned rubrics, answer keys and scoring guidelines that provide sufficient guidance for interpreting student performance. *A unit or longer lesson should:* ○ Use varied modes of curriculum-embedded assessments that may include pre-, formative, summative and self-assessment measures.
Rating: 3 2 1 0	**Rating: 3 2 1 0**	**Rating: 3 2 1 0**	**Rating: 3 2 1 0**

EQuIP Rubric for Lessons & Units: Mathematics

Directions: The Quality Review Rubric provides criteria to determine the quality and alignment of lessons and units to the Common Core State Standards (CCSS) in order to: (1) Identify exemplars/ models for teachers' use within and across states; (2) provide constructive criteria-based feedback to developers; and (3) review existing instructional materials to determine what revisions are needed.

Step 1 – Review Materials

- Record the grade and title of the lesson/unit on the recording form.
- Scan to see what the lesson/unit contains and how it is organized.
- Read key materials related to instruction, assessment and teacher guidance.
- Study and work the task that serves as the centerpiece for the lesson/unit, analyzing the content and mathematical practices the tasks require.

Step 2 – Apply Criteria in Dimension I: Alignment

- Identify the grade-level CCSS that the lesson/unit targets.
- Closely examine the materials through the "lens" of each criterion.
- Individually check each criterion for which clear and substantial evidence is found.
- Identify and record input on specific improvements that might be made to meet criteria or strengthen alignment.
- Enter your rating 0 – 3 for Dimension I: Alignment.

Note: Dimension I is non-negotiable. In order for the review to continue, a rating of 2 or 3 is required. If the review is discontinued, consider general feedback that might be given to developers/teachers regarding next steps.

Step 3 – Apply Criteria in Dimensions II – IV

- Closely examine the lesson/unit through the "lens" of each criterion.
- Record comments on criteria met, improvements needed and then rate 0 – 3.

When working in a group, individuals may choose to compare ratings after each dimension or delay conversation until each person has rated and recorded their input for the remaining Dimensions II – IV.

Step 4 – Apply an Overall Rating and Provide Summary Comments

- Review ratings for Dimensions I – IV adding/clarifying comments as needed.
- Write summary comments for your overall rating on your recording sheet.
- Total dimension ratings and record overall rating E, E/I, R, N – adjust as necessary.

If working in a group, individuals should record their overall rating prior to conversation.

Step 5 – Compare Overall Ratings and Determine Next Steps

- Note the evidence cited to arrive at final ratings, summary comments and similarities and differences among raters. Recommend next steps for the lesson/unit and provide recommendations for improvement and/or ratings to developers/teachers.

Additional Guidance on Dimension II: Shifts - When considering *Focus* it is important that lessons or units targeting additional and supporting clusters are sufficiently brief – this ensures that students will spend the strong majority of the year on major work of the grade. See the *K-8 Publishers Criteria for the Common Core State Standards in Mathematics*, particularly pages 8-9 for further information on the focus criterion with respect to major work of the grade at www.corestandards.org/assets/Math_Publishers_Criteria_K-8_Summer%202012_FINAL.pdf. With respect to *Coherence* it is important that the learning objectives are linked to CCSS cluster headings (see www.corestandards.org/Math).

Rating Scales

Rating for Dimension I: Alignment is non-negotiable and requires a rating of 2 or 3. If rating is 0 or 1 then the review does not continue.

Rating Scale for Dimensions I, II, III, IV:

3: Meets most to all of the criteria in the dimension
2: Meets many of the criteria in the dimension
1: Meets some of the criteria in the dimension
0: Does not meet the criteria in the dimension

Descriptors for Dimensions I, II, III, IV:

3: **Exemplifies CCSS Quality -** meets the standard described by criteria in the dimension, as explained in criterion-based observations.
2: **Approaching CCSS Quality -** meets many criteria but will benefit from revision in others, as suggested in criterion-based observations.
1: **Developing toward CCSS Quality -** needs significant revision, as suggested in criterion-based observations.
0: **Not representing CCSS Quality -** does not address the criteria in the dimension.

Overall Rating for the Lesson/Unit:

E: Exemplar – Aligned and meets most to all of the criteria in dimensions II, III, IV **(total 11 – 12)**
E/I: Exemplar *if* Improved – Aligned and needs some improvement in one or more dimensions **(total 8 – 10)**
R: Revision Needed – Aligned partially and needs significant revision in one or more dimensions **(total 3 – 7)**
N: Not Ready to Review – Not aligned and does not meet criteria **(total 0 – 2)**

Descriptor for Overall Ratings:

E: Exemplifies CCSS Quality – Aligned and exemplifies the quality standard and exemplifies most of the criteria across Dimensions II, III, IV of the rubric.
E/I: Approaching CCSS Quality – Aligned and exemplifies the quality standard in some dimensions but will benefit from some revision in others.
R: Developing toward CCSS Quality – Aligned partially and approaches the quality standard in some dimensions and needs significant revision in others.
N: Not representing CCSS Quality – Not aligned and does not address criteria.

EQuIP Task Review Rubric – Mathematics

Task Title_____ **Grade** _____ **Date**_____ **Rating:** ☐

Targeted Standards_____ **Mathematical Practices** _____

I. Alignment to the CCSS	II. Attention to the Instructional Shifts
The task clearly aligns with one or more CCSS. ☐ The performance expectations of the task address the mathematics, with precision and accuracy, for <u>at least part</u> of one CCSS. ☐ The task includes opportunities for a student to apply, and a teacher to observe, at least one Standard for Mathematical Practice.	*The task supports the key shifts that are reflected in the CCSS.* ☐ The task requires students to engage fully with the mathematics of the task, including providing opportunities for the appropriate aspects of rigor, <u>as required by the targeted standards</u>: • To independently apply mathematical concepts to real world situations • To apply their conceptual understanding of the mathematical content addressed • To practice and use core calculations and mathematical procedures quickly and accurately ☐ The task requires students to connect foundational knowledge to grade-level concepts, as required by the coherence in the standards. ☐ The task addresses, <u>or can be used to support</u>, a critical concept(s) for the grade level.
Notes and observations regarding alignment and attention to the instructional shifts of the CCSS:	

III. Support for Implementation

The task includes relevant supporting information or materials that ensure effective administration of the task and evaluation of student thinking.

☐ The task can be used to elicit direct, observable evidence of the degree to which each student can demonstrate the skills and knowledge addressed in the targeted CCSS.
☐ Supporting materials include answer keys, rubrics, and/or scoring guidelines that are clearly connected to the targeted CCSS and provide sufficient guidance for interpreting student performance.
☐ The task's prompts and directions provide sufficient guidance for the teacher to administer it effectively and for the students to complete it successfully.
☐ The task is accessible to and appropriate for <u>all</u> learners, including students who are English language learners or are working below or above grade level.
☐ The task cultivates student interest and/or engagement in the mathematics.

Notes and observations regarding support features that may be required for effective administration of the task:

Rating Descriptors:

E: Most criteria are checked, <u>including both in Dimension I</u> and those that are appropriate for the task's purpose in Dimension II and III. The task is likely to promote successful learning and/or assessment of the skills and knowledge required in the targeted CCSS.

E/I: Many criteria are checked. The task is aligned to the CCSS and has potential but could benefit from some minor improvements.

R: Some criteria are checked. The task has potential but needs significant revision to be considered effective.

N: The task is not recommended for instruction and/or assessment of the CCSS.

Assessment Evaluation Tool (AET)

Mathematics, Grades K–12

Assessment Evaluation Tool

Mathematics, Grades K–12

This Math AET is designed to help educators determine whether assessments and sets of assessments are aligned to the Shifts and major features of the Common Core State Standards (CCSS). The substantial instructional Shifts (http://www.corestandards.org/other-resources/key-shifts-in-mathematics/) at the heart of the Common Core State Standards in mathematics are:

- **Focus** strongly where the Standards focus
- **Coherence**: Think across grades and link to major topics within the grade
- **Rigor**: In major topics, pursue conceptual understanding, procedural skill and fluency, and application with equal intensity.

The AET draws directly from the following documents:

- Common Core State Standards for Mathematics (www.corestandards.org/Math)
- Publishers' Criteria for the Common Core State Standards for Mathematics, Grades K–8 (Spring 2013) (http://www.corestandards.org/wp-content/uploads/Math_Publishers_Criteria_K-8_Spring_2013_FINAL1.pdf), and Publishers' Criteria for the Common Core State Standards for Mathematics, High School (Spring 2013) (http://www.corestandards.org/wp-content/uploads/Math_Publishers_Criteria_HS_Spring_2013_FINAL1.pdf).

When to use the AET

1. Purchasing assessments: Many factors go into local purchasing decisions. Alignment to the Standards is a critical factor to consider. This tool is designed to evaluate alignment of assessments and sets of assessments to the Shifts and the major features of the CCSS. It also provides suggestions of additional indicators to consider in the assessment evaluation and purchasing process.
2. Evaluating assessments in use: The AET can be used to analyze the degree of alignment of existing assessments and sets of assessments and help to highlight specific, concrete flaws in alignment. Even where assessments currently in use fail to meet one or more of these criteria, the pattern of failure is likely to be informative. States and districts can use the evaluation to create a thoughtful plan to modify assessments and sets of assessments in such a way that they better meet the requirements of the Standards.
3. Developing assessments: This tool can be used to provide guidance for and evaluation of alignment for creating locally developed assessments and sets of assessments. States and districts creating new aligned assessments and sets of assessments should use the criteria within the AET to guide the development of test blueprints, item specifications, and item review.

Who Uses the AET

The AET is designed for use by educators and administrators including content specialists, assessment specialists, administrators and educators at the school, district or state level. The AET is designed for use by educators and administrators including content specialists, assessment specialists, administrators and educators at the school, district or state level. Evaluating assessments and sets of assessments requires both subject-matter and technical expertise. Evaluators should be well versed in the Standards (www.corestandards.org/Math) for all grades in which assessments are being evaluated. This includes understanding the Major Work of the grade (www.achievethecore.org/focus) and the widely applicable pre-requisites in high school (www.achievethecore.org/prerequisites), the Supporting and Additional work, how the content fits into the progressions in the Standards (www.achievethecore.org/progressions), and the expectations of the Standards with respect to conceptual understanding, procedural skill and fluency, and application. Evaluators also should be familiar with the substantial instructional Shifts (http://www.corestandards.org/other-resources/key-shifts-in-mathematics/) of Focus, Coherence and Rigor that are listed above.

Getting Started

Prior to Evaluation

Assemble all of the materials necessary for the evaluation, e.g., test blueprints, item specifications, operational forms, test items, metadata for those items, score reports, etc. It is essential for evaluators to have materials for all grades covered by the assessment program, as some criteria cannot be rated without having access to each grade. In addition, each evaluator should have a reference copy of the Common Core State Standards for Mathematics and the Publishers' Criteria for the Common Core State Standards for Mathematics, Grades K–8 (Spring 2013), and the Publishers' Criteria for the Common Core State Standards for Mathematics, High School (Spring 2013).

Sections 1–3 below should be completed to produce a comprehensive picture of the alignment to the Shifts and major features of the CCSSM for the assessments under evaluation. Information about areas in need of improvement should be shared with internal and external stakeholders.

Navigating the Tool

Begin with Section 1: Non-Negotiable Alignment Criteria (p. 4)

- The Non-Negotiable Alignment Criteria must each be met in full for assessments to be considered aligned to the Shifts and the major features of the Common Core State Standards. Each Non-Negotiable Alignment Criterion has one or more metrics associated with it; every one of these metrics must be met in order for the criterion as a whole to be met.
- Examine the relevant assessments and use evidence to rate the materials against each criterion and its associated metric(s).
- Record and explain the evidence upon which the rating is based.

Continue to Section 2: Alignment Criteria (p. 14)

- The Alignment Criteria must each be met for assessments to be considered aligned to the Shifts and major features of the Common Core State Standards. Each Alignment Criterion has one or more metric associated with it; a specific number of these metrics must be met or partially met in order for the criterion as a whole to be met.
- Examine the assessments in relation to these criteria, assigning each metric a point value. Rate the criterion as "Meets" or "Does Not Meet" based on the number of points assigned. The more points the assessments receive on the Alignment Criteria, the better they are aligned.
- Record and explain the evidence upon which the rating is based.

Complete Section 3: Evaluation Summary (p. 34)

- Compile all of the results from Sections 1 and 2 to determine if the assessments are aligned to the Shifts and major features of the CCSS.

Proceed to Section 4: Indicators of Quality (p. 36)

- Indicators of Quality are important considerations that will help evaluators better understand the overall quality of an assessment program. These considerations are not criteria for alignment to the CCSS, but they provide valuable information about additional program characteristics, such as ensuring accessibility for all students. Evaluators may want to add their own indicators to the examples provided.

Directions for Non-Negotiable 1

Focus on Major Work

Non-Negotiable 1: The large majority of points in each grade K–8 are devoted to the Major Work of the grade, and the majority of points in each high school course are devoted to widely applicable prerequisites.

Required Materials

- Test blueprints and operational forms
- "Focus by Grade Level" (achievethecore.org/focus) and the widely applicable prerequisites for postsecondary work (achievethecore.org/prerequisites).
- Publishers' Criteria for the Common Core State Standards for Mathematics, Grades K–8 (Spring 2013, pp. 8) (http://www.corestandards.org/wp-content/uploads/Math_Publishers_Criteria_K-8_Spring_2013_FINAL1.pdf)
- Publishers' Criteria for the Common Core State Standards for Mathematics, High School (Spring 2013, pp. 7) (http://www.corestandards.org/wp-content/uploads/Math_Publishers_Criteria_HS_Spring_2013_FINAL1.pdf)
- Common Core State Standards for Mathematics (corestandards.org/wp-content/uploads/Math_Standards.pdf)

Rating this Criterion

The metric will be rated as Meets or Does Not Meet/Insufficient Evidence. If the metric is rated as Does Not Meet/Insufficient Evidence, then the assessments fail this Non-Negotiable. If the metrics is rated as Meets, then the assessments pass this Non-Negotiable.

If the metric is rated as Meets, provide specific examples of evidence of this. If the assessment Does Not Meet the metric, include evidence of specific gaps found in the materials. If the materials provide Insufficient Evidence, explain what is missing from the materials or what within the materials is unclear.

Non-Negotiable 1

Focus on Major Work

Metric

NN Metric 1A:
For grades K–8, the assessment or set of assessments for each grade meet or exceed the following percentages:

- 85% or more of the total score points in the assessment(s) for each grade Kindergarten, 1, and 2 align exclusively to the Major Work of the grade.
- 75% or more of the total score points in the assessment(s) for each grade 3, 4, and 5 align exclusively to the Major Work of the grade.
- 65% or more of the total score points in the assessment(s) for each grade 6, 7, and 8 align exclusively to the Major Work of the grade.

For high school, the assessment or set of assessments for each course meet or exceed the following percentage:

50% or more of the total score points in each high school course assessment align to widely applicable prerequisites for postsecondary work.

Procedure for Evaluation

Familiarize yourself with the Major Work of the grade using the "Focus by Grade Level" documents and/or the widely applicable prerequisites using the "Widely Applicable Prerequisites" document.

Evaluate the blueprint or operational form(s) for each grade/course by counting the total number of points aligned to the Major Work of the grade or widely applicable prerequisites and divide by the total number of points on the test.

For context, read Criterion #1 in the Publishers' Criteria for the Common Core State Standards for Mathematics, Grades K–8 (Spring 2013) and Criterion #1 in the Publishers' Criteria for the Common Core State Standards for Mathematics, High School (Spring 2013).

Evidence

Rating

☐ **Meets**

☐ **Does Not Meet / Insufficient Evidence**

Non-Negotiable 1

Focus on Major Work

Non-Negotiable 1: The large majority of points in each grade K–8 are devoted to the Major Work of the grade, and the majority of points in each high school course are devoted to widely applicable prerequisites.

Rating for Non-Negotiable 1

If metrics were rated as Meets, then rate Non-Negotiable 1 as Meets. If one or more metrics were rated as Does Not Meet, then rate Non-Negotiable 1 as Does Not Meet. Check the final rating.

Then, briefly describe the strengths and weaknesses of these materials in light of this Criterion.

Rating

☐ **Meets**

☐ **Does Not Meet**

Strengths / Weaknesses:

Before moving to Non-Negotiable 2, record the final Meets or Does Not Meet rating in the Evaluation Summary on Page 34.

Directions for Non-Negotiable 2

Freedom from Major Obstacles to Focus

Non-Negotiable 2: No item assesses topics directly or indirectly before they are introduced in the CCSSM.

Required Materials

- Publishers' Criteria for the Common Core State Standards for Mathematics, Grades K–8 (Spring 2013, pp. 9) (http://www.corestandards.org/wp-content/uploads/Math_Publishers_Criteria_K-8_Spring_2013_FINAL1.pdf)
- Common Core State Standards for Mathematics (corestandards.org/wp-content/uploads/Math_Standards.pdf)
- Item specifications and operational forms or a representative sample of at least 20 operational items per grade/course
- "Focus by Grade Level" (achievethecore.org/focus) and the widely applicable prerequisites for postsecondary work (achievethecore.org/prerequisites).

Rating this Criterion

The metric will be rated as Meets or Does Not Meet/Insufficient Evidence. If the metric is rated as Does Not Meet/Insufficient Evidence, then the assessments fail this Non-Negotiable. If the metrics is rated as Meets, then the assessments pass this Non-Negotiable.

If the metric is rated as Meets, provide specific examples of evidence of this. If the assessment Does Not Meet the metric, include evidence of specific gaps found in the materials. If the materials provide Insufficient Evidence, explain what is missing from the materials or what within the materials is unclear.

Non-Negotiable 2

Freedom from Major Obstacles to Focus

Metric

NN Metric 2A:
100% of items on the assessment(s) assess knowledge of topics when they are introduced in the CCSSM.

Commonly misaligned topics include, but are not limited to:

- Probability, including chance, likely outcomes, probability models. (Introduced in the CCSSM in grade 7)
- Statistical distributions, including center, variation, clumping, outliers, mean, median, mode, range, quartiles; and statistical association or trends, including two-way tables, bivariate measurement data, scatter plots, trend line, line of best fit, correlation. (Introduced in the CCSSM in grades 6–8; see CCSSM for specific expectations by grade level.)
- Similarity, congruence, or geometric transformations. (Introduced in the CCSSM in grade 8)
- Symmetry of shapes, including line/reflection symmetry, rotational symmetry. (Introduced in the CCSSM in grade 4)

Procedure for Evaluation

Evaluate item specifications to see if content limits specify that the commonly misaligned topics listed in the metric are not assessed in grades prior to the grade introduced in the CCSSM.

Evaluate operational form(s) or a representative sample of at least 20 operational items per grade/course looking for commonly misaligned topics prior to the grade levels introduced by the CCSSM.

For context, read Criterion #2 in the Publishers' Criteria for the Common Core State Standards for Mathematics, Grades K–8 (Spring 2013).

Evidence

Rating

☐ **Meets**

☐ **Does Not Meet / Insufficient Evidence**

Non-Negotiable 2

Freedom from Major Obstacles to Focus

Non-Negotiable 2: No item assesses topics directly or indirectly before they are introduced in the CCSSM.

Rating for Non-Negotiable 2

If the metric was rated as Meets, then rate Non-Negotiable 2 as Meets. If metric was rated as Does Not Meet, then rate Non-Negotiable 2 as Does Not Meet. Check the final rating.

Then, briefly describe the strengths and weaknesses of these materials in light of this Criterion.

Rating

☐ **Meets**

☐ **Does Not Meet**

Strengths / Weaknesses:

Before moving to Non-Negotiable 3, record the final Meets or Does Not Meet rating in the Evaluation Summary on Page 34.

Directions for Non-Negotiable 3

Test Items Reflect the Coherence of the Standards

Non-Negotiable 3: Test items elicit direct, observable evidence of the degree to which a student can independently demonstrate the targeted Standard(s), reflecting the coherence of the CCSSM.

Required Materials

- Test blueprints and operational forms or a representative sample of at least 20 operational items per grade/course
- Metadata accompanying the items, showing the alignment of each question to the CCSS
- Publishers' Criteria for the Common Core State Standards for Mathematics, Grades K–8 (Spring 2013, pp. 13) (http://www.corestandards.org/wp-content/uploads/Math_Publishers_Criteria_K-8_Spring_2013_FINAL1.pdf)
- Publishers' Criteria for the Common Core State Standards for Mathematics, High School (Spring 2013, pp. 11 and 16) (http://www.corestandards.org/wp-content/uploads/Math_Publishers_Criteria_HS_Spring_2013_FINAL1.pdf)
- Common Core State Standards for Mathematics (http://corestandards.org/wp-content/uploads/Math_Standards.pdf)
- "Focus by Grade Level" (achievethecore.org/focus) and the widely applicable prerequisites for postsecondary work (achievethecore.org/prerequisites).

Rating this Criterion

Each metric will be rated as Meets or Does Not Meet/Insufficient Evidence. If any metric is rated as Does Not Meet/Insufficient Evidence, then the assessments fail this Non-Negotiable. If all metrics are rated as Meets, then the assessments pass this Non-Negotiable.

If the metric is rated as Meets, provide specific examples of evidence of this. If the assessment Does Not Meet the metric, include evidence of specific gaps found in the materials. If the materials provide Insufficient Evidence, explain what is missing from the materials or what within the materials is unclear.

Non-Negotiable 3

Test Items Reflect the Coherence of the Standards

Metric

NN Metric 3A:
Items exhibit alignment to the CCSSM for the grade or course by directly reflecting the language of individual Standards. All, or nearly all, items aligned to a single Standard should assess the central concern of the Standard in question.

Procedure for Evaluation

Evaluate operational form(s) or a representative sample of at least 20 operational items for each grade/course to check the alignment to the Standards for Mathematical Content. NOTE: An example of evaluating this metric might include ensuring that items aligned to 6.EE.A.3 put an emphasis on applying properties of operations and generating equivalent expressions, not just mechanically simplifying.

Evidence

Rating

☐ **Meets**

☐ **Does Not Meet / Insufficient Evidence**

Non-Negotiable 3

Test Items Reflect the Coherence of the Standards

Metric

NN Metric 3B:
Assessments exhibit alignment to the CCSSM for that grade or course: Operational forms for each grade/course include items that directly assess multiple levels of the content hierarchy (i.e. standard, cluster, and domain).

Procedure for Evaluation

Evaluate blueprints or operational form(s) for each grade/course to see if one or more items assess at the cluster, domain, or grade level.

For context, read Criterion #6 in the Publishers' Criteria for the Common Core State Standards for Mathematics, Grades K–8 (Spring 2013) and Criterion #4 in the Publishers' Criteria for the Common Core State Standards for Mathematics, High School (Spring 2013).

Evidence

Rating

☐ **Meets**

☐ **Does Not Meet / Insufficient Evidence**

Non-Negotiable 3

Test Items Reflect the Coherence of the Standards

Non-Negotiable 3: Test items elicit direct, observable evidence of the degree to which a student can independently demonstrate the targeted Standard(s), reflecting the coherence of the CCSSM.

Rating for Non-Negotiable 3

If metrics were rated as Meets, then rate Non-Negotiable 3 as Meets. If one or more metrics were rated as Does Not Meet, then rate Non-Negotiable 3 as Does Not Meet. Check the final rating.

Then, briefly describe the strengths and weaknesses of these materials in light of this Criterion.

Rating

☐ **Meets**

☐ **Does Not Meet**

Strengths / Weaknesses:

Before moving to Alignment Criterion 1, record the final Meets or Does Not Meet rating in the Evaluation Summary on Page 34.

Now continue by evaluating Alignment Criterion 1 for Rigor and Balance.

EQuIP Rubric for Lessons & Units: Science

Version 3.1

Introduction:

The Educators Evaluating the Quality of Instructional Products (EQuIP) Rubric for science provides criteria by which to measure the alignment and overall quality of lessons and units with respect to the Next Generation Science Standards (NGSS). The purposes of the rubric and review process are to: (1) review existing lessons and units to determine what revisions are needed; (2) provide constructive criterion-based feedback and suggestions for improvement to developers; (3) identify exemplars/models for teachers' use within and across states; and (4) to inform the development of new lessons and units.

To effectively apply this rubric, an understanding of the National Research Council's *A Framework for K–12 Science Education* and the *Next Generation Science Standards*, including the NGSS shifts (Appendix A of the NGSS), is needed. Unlike in the EQuIP Rubrics for mathematics and ELA, there is not a category in the science rubric for shifts. Over the course of the rubric development, writers and reviewers noted that the shifts fit naturally into the other three categories. For example, the blending of the three-dimensions, or three-dimensional learning, is addressed in each of the three categories; coherence is addressed in the first two categories; connections to the Common Core State Standards is addressed in the first category; etc. Each category includes criteria by which to evaluate the integration of engineering, when included in a lesson or unit, through practices or disciplinary core ideas. Another difference between the EQuIP Rubrics from mathematics and ELA is in the name of the categories; the rubric for science refers to them simply as *categories*, whereas the math and ELA rubrics refer to the categories as dimensions. This distinction was made because the Next Generation Science Standards already uses the term *dimensions* to refer to practices, disciplinary core ideas, and crosscutting concepts.

The architecture of the NGSS is significantly different from other sets of standards. The three dimensions, crafted into performance expectations, describe what is to be assessed following instruction and therefore are the measure of proficiency. A lesson or unit may provide opportunities for students to demonstrate performance of practices connected with their understanding of core ideas and crosscutting concepts as foundational pieces. This three-dimensional learning leads toward eventual mastery of performance expectations. In this scenario, quality materials should clearly describe or show how the lesson or unit works coherently with previous and following lessons or units to help build toward eventual mastery of performance expectations. The term *element* is used in the rubric to represent the relevant, bulleted practices, disciplinary core ideas, and crosscutting concepts that are articulated in the foundation boxes of the standards and in K–12 grade-banded progressions and the NGSS Appendices. Given the understanding that lessons and units should integrate the practices, disciplinary core ideas, and crosscutting concepts in ways that make sense instructionally and not replicate the exact integration in the performance expectations, the new term *elements* is needed to describe these smaller units of the three dimensions. Although it is unlikely that a single lesson would provide adequate opportunities for a student to demonstrate proficiency on an entire performance expectation, high-quality units are more likely to provide these opportunities to demonstrate proficiency on one or more performances expectations.

There is a recognition among educators that curriculum and instruction will need to shift with the adoption of the NGSS, but it is currently difficult to find instructional materials designed for the NGSS. The power of the rubric is in the feedback and suggestions for improvement it provides curriculum developers and the productive conversations in which educators engage while evaluating materials using the quality review process. For curriculum developers, the rubric and review process provide evidence of the quality and the degree to which the lesson or unit is designed for the NGSS. Additionally, the rubric and review process generate suggestions for improvement on how materials can be further improved and better designed to match up with the vison of the *Framework* and the NGSS.

Version 3.1 — published April 2021

EQuIP Rubric for Lessons & Units: Science

Lessons *and* units designed for the NGSS include clear and compelling evidence of the following*:*

I. NGSS 3D Design	II. NGSS Instructional Supports	III. Monitoring NGSS Student Progress
The lesson/unit is designed so students make sense of phenomena and/or design solutions to problems by engaging in student performances that integrate the three dimensions of the NGSS.	*The lesson/unit supports three-dimensional teaching and learning for ALL students by placing the lesson in a sequence of learning for all three dimensions and providing support for teachers to engage all students.*	*The lesson/unit supports monitoring student progress in all three dimensions of the NGSS as students make sense of phenomena and/or design solutions to problems.*
A. **Explaining Phenomena/Designing Solutions**: Making sense of phenomena and/or designing solutions to a problem drive student learning. i. Student questions and prior experiences related to the phenomenon or problem motivate sense-making and/or problem solving. ii. The focus of the lesson is to support students in making sense of phenomena and/or designing solutions to problems. iii. When engineering is a learning focus, it is integrated with developing disciplinary core ideas from physical, life, and/or earth and space sciences. B. **Three Dimensions**: Builds understanding of multiple grade-appropriate elements of the science and engineering practices (SEPs), disciplinary core ideas (DCIs), and crosscutting concepts (CCCs) that are deliberately selected to aid student sense-making of phenomena and/or designing of solutions. i. Provides opportunities to *develop and use* specific elements of the SEP(s). ii. Provides opportunities to *develop and use* specific elements of the DCI(s). iii. Provides opportunities to *develop and use* specific elements of the CCC(s). C. **Integrating the Three Dimensions:** Student sense-making of phenomena and/or designing of solutions requires student performances that integrate elements of the SEPs, CCCs, and DCIs.	A. **Relevance and Authenticity**: Engages students in authentic and meaningful scenarios that reflect the practice of science and engineering as experienced in the real world. i. Students experience phenomena or design problems as directly as possible (firsthand or through media representations). ii. Includes suggestions for how to connect instruction to the students' home, neighborhood, community and/or culture as appropriate. iii. Provides opportunities for students to connect their explanation of a phenomenon and/or their design solution to a problem to questions from their own experience. B. **Student Ideas:** Provides opportunities for students to express, clarify, justify, interpret, and represent their ideas and to respond to peer and teacher feedback orally and/or in written form as appropriate. C. **Building Progressions**: Identifies and builds on students' prior learning in all three dimensions, including providing the following support to teachers: i. Explicitly identifying prior student learning expected for all three dimensions ii. Clearly explaining how the prior learning will be built upon D. **Scientific Accuracy**: Uses scientifically accurate and grade-appropriate scientific information, phenomena, and representations to support students' three-dimensional learning. E. **Differentiated Instruction**: Provides guidance for teachers to support differentiated instruction by including: i. Supportive ways to access instruction, including appropriate linguistic, visual, and kinesthetic engagement opportunities that are essential for effective science and engineering learning and particularly beneficial for multilingual learners and students with disabilities. ii. Extra support (e.g., phenomena, representations, tasks) for students who are struggling to meet the targeted expectations. iii. Extensions for students with high interest or who have already met the performance expectations to develop deeper understanding of the practices, disciplinary core ideas, and crosscutting concepts.	A. **Monitoring 3D student performances**: Elicits direct, observable evidence of three-dimensional learning; students are using practices with core ideas and crosscutting concepts to make sense of phenomena and/or to design solutions. B. **Formative:** Embeds formative assessment processes throughout that evaluate student learning to inform instruction. C. **Scoring guidance**: Includes aligned rubrics and scoring guidelines that provide guidance for interpreting student performance along the three dimensions to support teachers in (a) planning instruction and (b) providing ongoing feedback to students. D. **Unbiased tasks/items**: Assesses student proficiency using methods, vocabulary, representations, and examples that are accessible and unbiased for all students.

EQuIP Rubric for Lessons & Units: Science

Units designed for the NGSS will *also include* clear and compelling evidence of the following additional criteria:

I. NGSS 3D Design	II. NGSS Instructional Supports	III. Monitoring NGSS Student Progress
D. **Unit Coherence**: Lessons fit together to target a set of performance expectations. i. Each lesson builds on prior lessons by addressing questions raised in those lessons, cultivating new questions that build on what students figured out, or cultivating new questions from related phenomena, problems, and prior student experiences. ii. The lessons help students develop toward proficiency in a targeted set of performance expectations. E. **Multiple Science Domains**: *When appropriate*, links are made across the science domains of life science, physical science and Earth and space science. i. Disciplinary core ideas from different disciplines are used together to explain phenomena. ii. The usefulness of crosscutting concepts to make sense of phenomena or design solutions to problems *across science domains* is highlighted. F. **Math and ELA:** Provides grade-appropriate connection(s) to the Common Core State Standards in Mathematics and/or English Language Arts & Literacy in History/Social Studies, Science and Technical Subjects.	F. **Teacher Support for Unit Coherence:** Supports teachers in facilitating coherent student learning experiences over time by: i. Providing strategies for linking student engagement across lessons (e.g. cultivating new student questions at the end of a lesson in a way that leads to future lessons, helping students connect related problems and phenomena across lessons, etc.). ii. Providing strategies for ensuring student sense-making and/or problem-solving is linked to learning in all three dimensions. G. **Scaffolded differentiation over time**: Provides supports to help students engage in the practices as needed and gradually adjusts supports over time so that students are increasingly responsible for making sense of phenomena and/or designing solutions to problems.	E. **Coherent Assessment system:** Includes pre-, formative, summative, and self-assessment measures that assess three-dimensional learning. F. **Opportunity to learn**: Provides multiple opportunities for students to demonstrate performance of practices connected with their understanding of disciplinary core ideas and crosscutting concepts and receive feedback.

Using the EQuIP Rubric for Lessons & Units: Science

The first step in the review process is to become familiar with the rubric, the lesson or unit, and the practices, disciplinary core ideas, and crosscutting concepts targeted in the lesson or unit. The three categories in the rubric are: NGSS 3D Design, NGSS Instructional Supports, and Monitoring NGSS Student Progress. Each criterion within each category should be considered separately as part of the complete review process and are used to provide sufficient information for determination of overall quality of the lesson or unit.

For the purposes of using the rubric, **a lesson is defined as**: a set of instructional activities and assessments that may extend over several class periods or days; it is more than a single activity. A **unit is defined as**: a set of lessons that extend over a longer period of time. If you are reviewing a lesson, you will use only the first section of the rubric (page 2). If you are reviewing an instructional unit, you apply all of the criteria of the rubric (pages 2 and 3) across the unit. You'll notice that the definition of a "unit" is intentionally broad here. If you are reviewing instructional materials that cover more than a few days of instruction, use the full unit list of criteria.

Also important to the review process is feedback and suggestions for improvement to the developer of the resource. For this purpose, a set of response forms is included so that the reviewer can effectively provide criterion-based feedback and suggestions for improvement for each category. The response forms correspond to the criteria of the rubric. Evidence for each criterion must be identified and documented and criterion-based feedback and suggestions for improvement should be given to help improve the lesson or unit.

While it is possible for the rubric to be applied by an individual, the quality review process works best with a team of reviewers, as a collaborative process, with the individuals recording their thoughts and then discussing with other team members before finalizing their feedback and suggestions for improvement. Discussions should focus on understanding all reviewers' interpretations of the criteria and the evidence they have found. With professional learning support for the group, this process will provide higher quality feedback about the lessons and also calibrate responses across reviewers in a way that moves them toward agreement about quality with respect to the NGSS. Commentary needs to be constructive, with all lessons or units considered "works in progress." Reviewers must be respectful of team members and the resource contributor. Contributors should see the review process as an opportunity to gather feedback and suggestions for improvement rather than to advocate for their work. All feedback and suggestions for improvement should be criterion-based and have supporting evidence from the lesson or unit cited.

In order to apply the rubric with reliability and with fidelity to its intent, it is recommended that those applying the rubric to lessons and units be supported to attend EQuIP professional learning based on the EQuIP Facilitator's Guide. There is guidance within the rubric below and in the Facilitator's Guide, but application of the rubric is much more successful with the support of professional learning. It is difficult to develop proficiency at using the rubric without *at least* two days of high quality professional learning that engages participants in evaluating lessons and units.

Step 1 – Review Materials

The first step in the review process is to become familiar with the rubric and the lesson or unit that is being evaluated.

- Review the rubric and record the grade and title of the lesson or unit on the response form.
- Scan the lesson/unit to see what it's about; identify what practices, disciplinary core ideas, and crosscutting concepts are targeted; and determine how it is organized.
- Read key materials related to instruction, assessment, and teacher guidance.
- Read the definitions of "lesson" and "unit" near the top of this page and decide as a group whether you will be using the shorter list of criteria for a lesson, or the longer list of criteria that apply to a unit.

Step 2 – Apply Criteria in Category I: NGSS 3D Design

Evaluate the lesson or unit using the criteria in the first category, first individually and then as a team.

- Closely examine the lesson or unit through the "lens" of each criterion in the first category.
- For each criterion, record where you find it in the lesson/unit (the evidence) and why/how this evidence is an indicator the criterion is being met (the reasoning)
- As individuals, check the box for each criterion on the response form that indicates the degree to which evidence could be identified.
- Identify and record input on specific improvements that might be made to meet criteria or strengthen alignment.

- Look across the criteria of the category (A–C for a lesson and A–F for a unit), evaluate the degree to which they are met, and enter your 0–3 rating for Category I: NGSS 3D Design (see scale description below)
- As a team, discuss criteria for which clear and substantial evidence is found, as well as criterion-based suggestions for specific improvements that might be needed to meet criteria. As a team, enter your 0–3 rating for Dimension I: NGSS 3D Design.

*If the rubric is being used to approve or vet resources and the lesson or unit does not score at least a "2" in **Category I: NGSS 3D Designed**, the review should stop and feedback should be provided to the lesson developer(s) to guide revisions. If the rubric is being used locally for revising and building lessons, professional judgment should guide whether to continue reviewing the lesson. Categories II and III may be time consuming to evaluate if Category I has not been met and the feedback may not be useful if significant revisions are needed in Category I, but evaluating these criteria in a group may support deeper and more common understanding of the criteria in these categories and more complete feedback to the lesson developer (if they are not in the room) so that Categories II and III are more likely to be met with fewer cycles of revision.*

Step 3 – Apply Criteria in Categories II and III: Instructional Supports and Monitoring Student Progress

The third step is to evaluate the lesson or unit using the criteria in the second and third categories, first individually and then as a group.

- Closely examine the lesson or unit through the "lens" of each criterion in the second and third categories of the response form.
- For each criterion, record where you find it in the lesson/unit (the evidence) and why/how this evidence is an indicator the criterion is being met (the reasoning)
- Individually check the box for each criterion on the response form that indicates the degree to which evidence could be identified.
- Record any suggestions for improvement and then rate each category using the 0–3 rating scale in the forms below.

When working in a group, teams may choose to compare ratings after each category or delay conversation until each person has rated and recorded input for both Categories II and III. Complete consensus among team members is not required but discussion is a key component of the review process that moves the group to a better understanding of the criteria.

Step 4 – Apply an Overall Rating and Provide Summary Comments

- Review ratings for Categories I–III, adding/clarifying comments as needed.
- Write summary comments for your overall rating on your recording sheet.
- Total category ratings, reflect on the overall quality of the lesson or unit, and record the overall rating of E, E/I, R, or N.

If working in a group, individuals should record their overall rating prior to conversation.

Step 5 – Compare Overall Ratings and Recommend Next Steps

- Note the evidence cited to arrive at final ratings, summary comments and similarities and differences among raters. Recommend next steps for the lesson/unit and provide recommendations for improvement and/or ratings to developers/teachers.

Rating Scales

Rating for Category I: NGSS 3D Designed is non-negotiable and requires a rating of 2 or 3. If rating is 0 or 1 then a review for resource approval does not continue.

Rating Scale for Categories I, II, & III:

Rating scales are different for each category and can be found after each category in the rubric.

Descriptors for Categories I, II, & III:

3: **Exemplifies NGSS Quality**—meets the standard described by criteria in the category, as explained in criterion-based observations.
2: **Approaching NGSS Quality**—meets many criteria but will benefit from revision in others, as suggested in criterion-based observations.
1: **Developing toward NGSS Quality**—needs significant revision, as suggested in criterion-based observations.
0: **Not representing NGSS Quality**—does not address the criteria in the category.

Overall Rating for the Lesson/Unit:

E: Example of high quality NGSS design—High quality design for the NGSS across all three categories of the rubric; a lesson or unit with this rating will still need adjustments for a specific classroom, but the support is there to make this possible; exemplifies most criteria across Categories I, II, & III of the rubric. (total score ~8–9)
E/I: Example of high quality NGSS design if Improved—Adequate design for the NGSS, but would benefit from some improvement in one or more categories; most criteria have at least adequate evidence (total score ~6–7)
R: Revision needed—Partially designed for the NGSS, but needs significant revision in one or more categories (total ~3–5)
N: Not ready to review—Not designed for the NGSS; does not meet criteria (total 0–2)

Reviewer Name or ID: _____ **Grade:**_____ **Lesson/Unit Title:**_____

Category I: NGSS 3D Design (lessons and units): *The lesson/unit is designed so students make sense of phenomena and/or design solutions to problems by engaging in student performances that integrate the three dimensions of the NGSS.*

Lesson *and* Unit Criteria Lessons and units designed for the NGSS include clear and compelling evidence of the following:	**Specific evidence from materials** (what happened/where did it happen) **and reviewer's reasoning** (how/why is this evidence)		**Evidence of Quality?**	**Suggestions for improvement**
A. **Explaining Phenomena/Designing Solutions**: Making sense of phenomena and/or designing solutions to a problem drive student learning. i. Student questions and prior experiences related to the phenomenon or problem motivate sense-making and/or problem solving. ii. The focus of the lesson is to support students in making sense of phenomena and/or designing solutions to problems. iii. When engineering is a learning focus, it is integrated with developing disciplinary core ideas from physical, life, and/or earth and space sciences.			☐ None ☐ Inadequate ☐ Adequate ☐ Extensive	
B. **Three Dimensions**: Builds understanding of multiple grade-appropriate elements of the science and engineering practices (SEPs), disciplinary core ideas (DCIs), and crosscutting concepts (CCCs) *that are deliberately selected to aid student sense-making of phenomena and/or designing of solutions*.	Document evidence and reasoning, and evaluate whether or not there is sufficient evidence of quality for each dimension separately	**Evidence of Quality?**	☐ None ☐ Inadequate ☐ Adequate ☐ Extensive **(All 3 dimensions must be rated at least "adequate" to mark "adequate" overall)**	
i. Provides opportunities to *develop and use* specific elements of the SEP(s).	i.	☐ None ☐ Inadequate ☐ Adequate ☐ Extensive		
ii. Provides opportunities to *develop and use* specific elements of the DCI(s).	ii.	☐ None ☐ Inadequate ☐ Adequate ☐ Extensive		
iii. Provides opportunities to *develop and use* specific elements of the CCC(s). **Evidence needs to be at the *element level* of the dimensions (see rubric introduction for a description of what is meant by "element")**	iii.	☐ None ☐ Inadequate ☐ Adequate ☐ Extensive		

C. **Integrating the Three Dimensions:** Student sense-making of phenomena and/or designing of solutions requires student performances that integrate elements of the SEPs, CCCs, and DCIs.		☐ None ☐ Inadequate ☐ Adequate ☐ Extensive	
Rating for Category I. NGSS 3D Design—*lessons* After carefully weighing the evidence, reasoning, and suggestions for improvement, rate the degree to which there is enough evidence to support a claim that the lesson meets these criteria. ***If you are evaluating an instructional unit rather than a single lesson, continue on to evaluate criteria D-F and rate Category I overall below.***	***Lesson* Rating scale for Category I (Criteria A–C only)**: **3**: Extensive evidence to meet at least two criteria (and at least adequate evidence for the third) **2**: Adequate evidence to meet all three criteria in the category **1**: Adequate evidence to meet at least one criterion in the category, but insufficient evidence for at least one other criterion **0**: Inadequate (or no) evidence to meet any of the criteria in the category		**Circle Rating** **0 1 2 3** After rating the lesson, read below for next steps

What's next if the lesson rating is less than a 2?

*If the rubric is being used to approve or vet resources and the lesson or unit does not score at least a "2" in **Category I: NGSS 3D Designed**, the review should stop and feedback should be provided to the lesson developer(s) to guide revisions. If the rubric is being used locally for revising and building lessons, professional judgment should guide whether to continue reviewing the lesson. Categories II and III may be time consuming to evaluate if Category I has not been met and the feedback may not be useful if significant revisions are needed in Category I, but evaluating these criteria in a group may support deeper and more common understanding of the criteria in these categories and more complete feedback to the lesson developer (if they are not in the room) so that Categories II and III are more likely to be met with fewer cycles of revision.*

What's next if the lesson rating is a 2 or 3?

If you are evaluating a lesson that shows sufficient evidence of quality to warrant a rating of either a 2 or a 3 for Category I, proceed to Category II: NGSS Instructional Supports

Category I: NGSS 3D Design (additional criteria *for units only*):

If you are evaluating a lesson, it is not necessary to evaluate criteria D–F. Please enter your rating for a single lesson above (after C).

Unit Criteria A unit or longer lesson designed for the NGSS will also include clear and compelling evidence of the following:	**Specific evidence from materials and reviewers' reasoning**	**Evidence of Quality?**	**Suggestions for improvement**
D. **Unit Coherence**: Lessons fit together to target a set of performance expectations. i. Each lesson builds on prior lessons by addressing questions raised in those lessons, cultivating new questions that build on what students figured out, or cultivating new questions from related phenomena, problems, and prior student experiences. ii. The lessons help students develop toward proficiency in a targeted set of performance expectations.		☐ None ☐ Inadequate ☐ Adequate ☐ Extensive	
E. **Multiple Science Domains**: *When appropriate*, links are made across the science domains of life science, physical science and Earth and space science. i. Disciplinary core ideas from different disciplines are used together to explain phenomena. ii. The usefulness of crosscutting concepts to make sense of phenomena or design solutions to problems *across science domains* is highlighted.		☐ None ☐ Inadequate ☐ Adequate ☐ Extensive	
F. **Math and ELA:** Provides grade-appropriate connection(s) to the Common Core State Standards in Mathematics and/or English Language Arts & Literacy in History/Social Studies, Science and Technical Subjects.		☐ None ☐ Inadequate ☐ Adequate ☐ Extensive	
Rating for Category I. NGSS 3D Designed—*units* After carefully weighing the evidence, reasoning, and suggestions for improvement, rate the degree to which the criteria are met across the unit.	**Unit Rating Scale for Category I (Criteria A–F):** **3**: At least adequate evidence for all of the unit criteria in the category; extensive evidence for criteria A–C **2**: At least some evidence for all unit criteria in Category I (A–F); adequate evidence for criteria A–C **1**: Adequate evidence for some criteria in Category I, but inadequate/no evidence for at least one criterion A–C **0**: Inadequate (or no) evidence to meet any criteria in Category I (A–F)		**Circle Rating** **0 1 2 3**

*If the rubric is being used to approve or vet resources and the unit does not score at least a "2" overall in **Category I: NGSS 3D Design**, the review should stop here and feedback should be provided to the unit developer(s) to guide revisions. If the rubric is being used locally for revising and building units, professional judgment should be used on whether or not to continue reviewing the unit. For example, a unit that is weak in one aspect of criterion A, but that the reviewers think is easy to fix, might warrant continued review to provide more complete feedback to the unit developer(s).*

Science Task Prescreen

Introduction

The purpose of the Science Task Prescreen is to conduct a quick review of assessment tasks to determine whether they might be designed for standards based on the *Framework for K-12 Science Education*, like the Next Generation Science Standards (NGSS). The Prescreen is intended to reveal whether tasks include "red flags"—i.e., challenges commonly found in science assessment tasks.

Evaluating tasks using the Prescreen questions can help educators decide whether a task is worth diving into more deeply. Those interested in pursuing a more rigorous evaluation of tasks should use the Science Task Screener; however, the Task Screener assumes a deeper understanding of *A Framework for K–12 Science Education* and the NGSS. Those who are familiar with the assessment tasks, but not very familiar with the *Framework* or NGSS, should start with the Prescreen as a bridge to understand the key features of tasks developed for the NGSS and *Framework*. For those less familiar with the *Framework*, it will be particularly helpful to use the Prescreen as part of a collaborative professional learning process, to help build a common understanding of the questions and what constitutes as evidence to address them.

Because the Prescreen is a quick screening tool as opposed to a comprehensive evaluation tool, the questions in the Prescreen focus on features that are non-negotiable, easily identified, and reflect the most serious "fatal flaws" seen in attempts to develop science tasks. While there are indeed many other critically important features of science assessments, they are excluded here for the purposes of screening, and are addressed in the Task Screener. For more information about how the Prescreen was developed and its relationship to the Task Screener, please see these Frequently Asked Questions .

Using the Task Prescreen to evaluate science assessment tasks

While it is possible for the Prescreen to be applied by an individual, it is more powerful when used as part of a collaborative review process. These high-level questions can drive very meaningful conversations and help reviewers come to a common understanding of features of NGSS tasks. Reviewers should carefully discuss their answers to the questions and the evidence in the task that led them to those answers to come to a common understanding of language and expectations.

The Prescreen is organized around a short series or yes-or-no questions. In applying the Prescreen to a task, follow these simple steps:

1. Read through the task and complete the task as though you were a student.
2. Read through any additional support materials for the task.
3. Answer the questions in the Prescreen regarding the task and note any red flags.
4. Discuss the answers to the questions and evidence to support those answers with other reviewers.
5. Use your analysis to determine next steps for the task.

Because the Prescreen is applied at the level of the task rather than individual questions, reviewers will need to answer the questions based on evidence from the task as a whole. After reviewing the task using the Prescreen, reviewers should consider the red flags they have identified and determine, based on their needs, whether the assessment:

A. **Warrants further review.** If tasks have few red flags, they might be effective tasks and would benefit from a deeper evaluation. This might be particularly relevant for assessments that are used as major components of a lesson or unit; used across multiple classrooms or schools; or used in other high-impact, higher-stakes scenarios, such as tasks used as part of district- or state-wide assessment efforts. Red flags can be used to determine if the assessment has potential and to focus the major areas of improvement that might be needed.

B. **Should not be used.** Reviewers can use the red flags to determine that, for their current purposes, the task should simply not be used.

Science Task Prescreen

Task Title_____ **Grade** _____ **Date** _____

SEP: _____ **DCI:**_____ **CCC:** _____

Task Purpose:_____

Before you begin: Complete the task as a student would. Then, consider any support materials provided to teachers or students, such as contextual information about the task and answer keys/rubrics.

Prescreen: Answer the following high-level questions to identify any major red flags (🚩) in your task. If you find one or more red flags, consider the purpose of the task and the evidence gathered to determine whether the task warrants a deeper dive.

Question	Yes	No
1. Is there a **phenomenon or problem driving the task?**	☐	☐ 🚩
2. Can the majority of the task be answered **without** using information provided by the task scenario?	☐ 🚩	☐
3. Can significant portions of the task be answered successfully by using **rote knowledge** (e.g., definitions, prescriptive or memorized procedure)?	☐ 🚩	☐
4. Does the majority of the task require students to **use reasoning** to successfully complete the task?	☐	☐ 🚩
5. Does the task require students to use some understanding of **disciplinary core ideas** to successfully complete the task?	☐	☐ 🚩
6. Do students have to use at least one **science and engineering practice** to successfully complete the task?	☐	☐ 🚩
7. Are the **dimensions assessed separately** in the majority of the task?	☐ 🚩	☐
8. Is the task **coherent and comprehensible** from the student perspective?	☐	☐ 🚩

Based on your assessment needs and the task purpose recorded above, make a recommendation about this task moving forward (choose one) :

☐ Warrants further review.

☐ Should not be used.

Summarize your evidence and reasoning:

CRITERIA FOR PROCURING AND EVALUATING HIGH-QUALITY AND ALIGNED SUMMATIVE SCIENCE ASSESSMENTS

Version 1.0 – March 2018

I. IN TRODUCTION

A growing number of states have demonstrated a commitment to ensuring better outcomes for all students by developing, adopting, and implementing rigorous science standards based on the National Research Council's *A Framework for K-12 Science Education*, such as the Next Generation Science Standards (NGSS). Fully meeting the vision set forth by the *Framework* and standards designed to implement it requires high-quality and aligned assessments that can provide actionable information to students, teachers, and families. Three-dimensional standards—those that integrate the Science and Engineering Practices (SEPs), Crosscutting Concepts (CCCs), and Disciplinary Core Ideas (DCIs)—based on the *Framework* are comprehensive, and it is unlikely that most states will assess the full range of depth and breadth in a single summative assessment opportunity for each student. States have several decisions to make regarding how to translate the depth and breadth of their science standards into appropriate statewide summative science assessments. While those decisions will vary from state to state, there *is* a common vision underlying all three-dimensional assessment efforts—and this document describes the criteria that define those common features in a statewide summative assessment.

Achieve developed this document with extensive input from experts and practitioners in the science and assessment fields. It is grounded in our collective and evolving understanding of how best to assess multi-dimensional standards, in the research that defines what all students should know and be able to do in science, and in lessons learned from early state processes in developing three-dimensional assessments. Regardless of each state's approach, this document is intended to be a useful resource for anyone developing and/or evaluating statewide summative assessments aligned to their *Framework*-based three-dimensional science standards.

THE PURPOSE AND AUDIENCE FOR THIS DOCUMENT

This document describes the features of a statewide summative science assessment that has been designed to embody standards based on the *Framework for K-12 Science Education*, such as the NGSS—to reflect its intent, grounded in the specific expectations of three-dimensional standards. Importantly, this document outlines the expectations for high-quality statewide summative science assessments that are designed and administered, in part, to meet federal requirements for science testing under Title I Part A of the Every Student Succeeds Act. As such, the criteria and evidence described here are grounded in the expectations outlined in the *Framework* and the NGSS as well as those described by federal peer review guidelines. In other words, while the priority for these criteria is to embody the intent of the NGSS and *Framework,* they are intentionally bounded by what would be needed and feasible to meet federal expectations for statewide summative assessments. They do not describe the expectations for other forms of science assessments that states and districts might use, such as interim or benchmark assessments or classroom-embedded summative and formative assessments. As such, expectations for a complete state system of science assessment is beyond the scope of this document. It is important to note that this is not because specifying the criteria for a full system of assessments is not important, but because this is a common component of the assessment system that all states are grappling with.

This document is intended to support state assessment directors, science supervisors, science assessment leads, test developers, and organizations that conduct independent evaluations of alignment of statewide summative assessments to state standards.

TERMINOLOGY USED IN THIS DOCUMENT.

Throughout this document, the term 'assessment' is used to refer to the full suite of statewide summative science assessments being developed or selected by a state for a given grade level (inclusive of multiple forms, years of administration, etc.). Some of the evidence descriptors are specific to what an evaluator might examine on an operational test form (the tests that students might see, plus answer keys and associated alignment claims)—these are labeled as 'test forms' and are distinguished from 'documentation', which include supporting information that relates to the development and interpretation of the entire assessment suite.

The term 'tasks' is used instead of the more traditional 'items' to better reflect the nature of questions on assessments designed for Framework-based standards. A task includes all scenario/stimuli and prompts associated with a common activity; it can utilize multiple item formats, can have multiple parts, and can require students to respond to open-ended questions. The term 'prompt' is used to identify the specific questions associated with a task. Generally, one or more prompts combine to form a task. A 'scenario' is the phenomenon- or problem-based context used to engage students in the scientific thinking required by the task. A scenario is coherent, engaging, relevant, and provides students with the scientific information (descriptions, data, models, arguments, etc.) they need to successfully respond to the task using the SEPs, CCCs, and DCIs targeted by the task. Throughout the document, 'targeted standards' are referenced—these indicate the state standards a task is intended to assess, and includes both complete performance expectations as well as the specific SEPs, CCCs, and DCIs.

This document contains science-specific (e.g., scientific interpretations of the word 'evidence') and NGSS-specific (e.g., the use of the word 'element' to refer to the specific bulleted ideas described in the *Framework* and the NGSS appendices) uses of words and phrases to convey intentional ideas. A full glossary of specific language uses can be found in Appendix A.

This document is also built on some key principles underlying assessments for which these criteria are appropriate. These principles are detailed in Appendix B.

EQUITY IN SCIENCE ASSESSMENTS

Ensuring that all students, including those from non-dominant groups, have access to a high-quality and rigorous science education that prepares them for college, career, and citizenship is at the heart of the *Framework* and the NGSS. This emphasis on student equity must extend to current efforts in assessments. Because statewide summative assessment data is used to evaluate and act on student science proficiency among student subgroups, it is imperative that *Framework*-based tests intentionally support students from non-dominant communities in demonstrating their scientific knowledge and abilities. It is difficult to make a validity argument for an assessment if students are incorrectly answering questions because of linguistic barriers or language mismatch, poor engagement, cultural insensitivities or bias, or inappropriately signaled scenarios that lead

students to answer the posed questions without using the targeted knowledge and skill. Because other resources provide extensive guidance about general accessibility and accommodations in assessments, this document focuses on the aspects of student equity and diversity that are most closely tied to content on science assessments, including the design of phenomena, problems, and tasks eliciting three-dimensional performances from students. This is embedded throughout the criteria, rather than posed as a separate expectation, to emphasize that a focus on equity cannot be separated from expectations for high-quality and aligned assessments—one cannot have a high-quality assessment that doesn't support all students. For more detail about how diversity and equity are included in each criterion, please see the FAQs.

II. OVERVIEW OF SCIENCE ALIGNMENT CRITERIA

The criteria for science build on those described for mathematics, English language arts, and testing practice by the CCSSO Criteria for Procuring and Evaluating High Quality Assessments (CCSSO, 2014). Like the CCSSO Criteria for aligned mathematics and ELA assessments, the current document describes the features all science assessments should demonstrate to be considered aligned to *Framework*-based science standards, as well as the kinds of evidence test developers could provide to show how well a given assessment meets the criteria. These criteria and associated evidence descriptors describe the baseline of common features for assessments. As states articulate their goals and intended uses for their science assessment, they may add to the criteria as appropriate. Additionally, the criteria challenge states to envision three-dimensional items, which are accessible by all students and grounded in the vision of the *Framework for K-12 Science Education*.

To demonstrate it is aligned to the NGSS or similar Framework-based standards, statewide summative science assessments must meet the following criteria:

Criterion	Description
1. **Design.** Assessments are intentionally designed to assess state science standards in order to provide evidence to support, refute, or qualify state-specific claims about students' achievement in science.	Assessment tasks, and the precise determinations of how well they align to standards, are informed by the design of the assessment, including how tasks individually and collectively provide valid evidence to support an assessment's claims and reporting priorities, and under what conditions.
2. **Three-dimensional performance.** Assessments require students to make sense of phenomena and solve problems by integrating the three dimensions. Assessment tasks elicit sense-making and problem solving by focusing strongly on reasoning using scientific and engineering evidence, models, and principles.	Assessments provide evidence of student knowledge and practice described by the targeted standards by requiring students to use the three dimensions (SEPs, CCCs, and DCIs) to identify and interpret evidence and engage in scientific reasoning as they make sense of phenomena and address problems.
3. Phenomena. Assessment scenarios focus on relevant, engaging, and rich phenomena and problems that elicit meaningful student performances. Assessment tasks are driven by meaningful and engaging scenarios.	Assessment tasks are situated in the context of meaningful scenarios, and are designed to elicit grade-appropriate, three-dimensional responses (i.e., responses in which students use multiple dimensions together).

4. **Scope.** Assessments are balanced across domains, and assess a range of knowledge and application within each dimension.	The summative assessments sample across conceptual understanding of core science ideas and crosscutting concepts, elements of scientific practices, and purposeful application of science as described by *Framework*-based standards.
5. Cognitive complexity. Assessments require a range of analytical thinking.	The assessments allow for robust information to be gathered for students with varied levels of achievement by providing opportunities that require all students to demonstrate varying levels of reasoning across life, physical, and Earth and space sciences as well as engineering, via SEPs and CCCs that range in grade-appropriate sophistication. Accommodations maintain the range of higher order analytical thinking skills as appropriate.
6. Technical Quality. Assessment tasks are of high technical quality and represent varied task types.	High-quality, fair, and unbiased tasks and a variety of types are strategically used to assess the standard(s). Tasks are designed with a focus on ensuring students from non-dominant communities are supported in demonstrating what they know and can do in science.
7. Reports. Assessments reports yield valuable information on student progress toward three-dimensional learning**.**	Assessment reports should be designed with specific uses in mind, transparently detail those uses, and illustrate student progress on the continuum toward the goals established by the standards at each grade band. Reports . Reports should focus on connecting the assessment purpose and appropriate uses of the assessment information, and on the integration and application of the knowledge and abilities described by the standards, and how they are addressed by the assessment..

This document does not address every aspect of assessment design that would need to be considered as states develop and evaluate their assessments; rather, it focuses on the features of content alignment (across all three dimensions) to the *Framework* and the NGSS. Many of the other important considerations states will have to contend with (e.g., accessibility) are addressed in the CCSSO Criteria.

The criteria, and evidence needed to meet the criteria, presented in this document represent a few notable shifts from traditional alignment expectations:

1) **The importance of an intentional design approach**. Traditional conceptualizations of alignment, that prioritize how well items "hit" targeted standards and "cover" the breadth of standards will not work for the NGSS given the breadth and depth of expectations both within a given standard and across the range of standards for a given grade level or band. To effectively assess the NGSS within common summative testing constraints, states will need to establish their priorities for the assessment. For example, states will need to determine:

Washington Quality Review Rubric for Social Studies Lessons & Units

The purpose of this rubric is to measure the alignment and overall quality of lessons and units with respect to the Washington State Social Studies Learning Standards and the Washington ELA and Literacy in History/Social Studies Standards. This rubric also evaluates lessons and units for integration with the College, Career, and Civic Life (C3) Framework. The rubric is divided into four dimensions:

I. Alignment to Standards
II. Teaching Strategies
III. Instructional Supports
IV. Assessment

Each of these dimensions is rated on a scale of 0 to 3:
0—Does not meet any of the criteria in the dimension
1—Meets some of the criteria in the dimension
2—Meets many of the criteria in the dimension
3—Meets all of the criteria in the dimension

This rubric is designed to evaluate:

Lessons that include instructional activities and assessments that may extend over a few class periods or days
Units that include integrated and focused lessons that extend over a longer period of time

The rubric is not designed to evaluate a single task or stand-alone activity.

Intended Use:

Review existing lessons and units to determine what revisions or supplements are needed

Inform the development of new lessons and units

Build the capacity of educators to evaluate and improve the quality of instructional materials for use in their classrooms and schools.

Reviewed resources may be in either print, digital, or online formats. They may carry different licensing types from open educational resources (OER) to all rights reserved.

Review Process

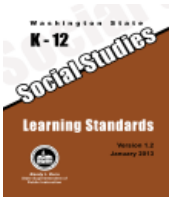
- Reviews using this rubric are best accomplished collaboratively, with team members providing specific evidence of how a resource meets dimension criteria and discussing the results.
- Look at the criteria in each dimension through the lens of the intended grade band.
- Check a criterion box only if there is clear and substantial evidence of the criterion (there are no "half-checks"). There may be instances when reviewers find clear and substantial evidence of a criterion and there are still constructive suggestions that can be made. In such cases, reviewers may provide feedback related to criteria that have been checked.
- For some resources, certain criteria will not be applicable. As a result, it's acceptable to give a "3" rating without having all of the criteria checked within a dimension; just support all ratings with specific evidence.

If recommendations for improvement are too significant, then the rating should be less than a "3." There should be a relationship between the number of checks and the overall rating. There shouldn't be huge misalignment, but it comes down to professional judgment. Reviewers should stand back and look at the review in its totality.

Washington Quality Review Rubric for Social Studies Lessons & Units – Version 2.3

Recommendations:

To effectively apply this rubric, an understanding of the Washington State Social Studies Learning Standards (GLEs), the Washington State ELA and Literacy in History/Social Studies Standards (Common Core State Standards), and the C3 Framework for Social Studies State Standards is needed.



Notes

- Materials from other states may need to be adapted to work within Washington learning standards. Prior to any adaptation, make sure edits are permitted under the resource license type.
- Even an exemplary unit may have to be adapted to meet the specific needs of your learners. Teaching is a changing practice; there will always be new and updated resources.

Additional Resources

Office of Superintendent of Public Instruction - Social Studies
Washington State Social Studies Laws/Regulations
Washington Social Studies Teachers Connect
Washington State Council for the Social Studies
C3 Resources from the C3 Literacy Collaborative

We express our gratitude to all the educators involved in the adaptation of this rubric. Without their support and expertise in the field of Social Studies, this resource would not be possible. This work was funded through a grant from the Washington State OER Project and administered by Educational Service District 105.

version 2.3
updated 5/31/2016

Unit Title: _____ **Reviewer ID:** _____

I. Alignment to Standards
The lesson/unit: ☐ Targets a set of grade-level standards in the Washington State Social Studies Learning Standards (GLEs) in one or more of the following areas: Geography, Civics, Economics, History, or Social Studies Skills. ☐ Integrates social studies content knowledge with reading, writing, speaking, and listening skills as outlined in the Washington State ELA and Literacy in History/Social Studies Standards (Common Core State Standards) *K–5: pages 9–33* *6-12: pages 59–66*
Summary of Observations and Suggestions for Improvement:
Rating: 3 2 1 0
Rating Scale for Dimensions I, II, III, IV: 3: Meets most to all of the criteria in the dimension. 2: Meets many of the criteria in the dimension. 1: Meets some of the criteria in the dimension. 0: Does not meet the criteria in the dimension.

Unit Title: _____ **Reviewer ID:** _____

II. Teaching Strategies

The lesson/unit infuses the strategies in the C3 Framework for Social Studies State Standards.

- ☐ **Integrates content and skills purposefully:** Thoughtfully introduces appropriate and relevant content for students to ground their inquiries and build disciplinary skills and conceptual knowledge.
- ☐ **Crafts questions that spark and sustain inquiry:** Provides deeper-level questions and/or gives students the opportunity to construct compelling and supporting questions to initiate and sustain an inquiry.
- ☐ **Students work collaboratively**: Engages students in disciplinary content to develop, examine, and communicate ideas.

The lesson/unit promotes literacy practices in the Washington State Learning Standards.

- ☐ **Credible primary and/or secondary sources**: When applicable, resource includes multiple perspectives.
- ☐ **Grade-level texts**: Resource includes readings that match grade band text complexity and are of sufficient quality and scope for the stated purpose.
- ☐ **Text-based evidence:** Facilitates rich and rigorous evidence-based discussions and writing through specific, thought-provoking questions.
- ☐ **Writing from sources:** Routinely expects that students draw and properly cite evidence from texts to inform, explain, or make an argument in a written form (notes, summaries, short responses or formal essays).
- ☐ **Academic vocabulary:** Focuses on building students' academic vocabulary in context throughout instruction.
- ☐ **Research:** Builds and presents knowledge through the process of analysis and synthesis as appropriate.

A longer lesson or unit should also:

- ☐ Increase text complexity: Focuses students on reading a progression of complex texts where the learning is sequenced, scaffolded, and supported to advance students toward independent reading.
- ☐ Make reading text closely and examining textual evidence a factor of the instructional focus.
- ☐ Build disciplinary knowledge in one or more of the following Washington State social studies strands (civics, economics, geography, history, and social studies skills).
- ☐ Provide state, tribal, and other perspectives, when applicable, while presenting or contrasting the unit within a global context.
- ☐ Integrate 21st Century skills (creativity and innovation, critical thinking and problem solving, communication, and collaboration).
- ☐ Provide tangible opportunities for taking informed action: Students, where curricularly appropriate, have the opportunity to culminate their academic inquiries through civic engagement.

Summary of Observations and Suggestions for Improvement:

Rating: 3 2 1 0

Unit Title: _____ **Reviewer ID:** _____

Rating Scale for Dimensions I, II, III, IV:
3: Meets most to all of the criteria in the dimension.
2: Meets many of the criteria in the dimension.
1: Meets some of the criteria in the dimension.
0: Does not meet the criteria in the dimension.

.

Unit Title: _____ **Reviewer ID:** _____

III. Instructional Supports

The lesson/unit is responsive to varied student learning needs.

- ❒ Includes a clear and specific purpose for instruction as well as specific guidance to support teaching and learning of targeted standards.
- ❒ Cultivates student interest and engagement in history/social studies.
- ❒ Supports learning of the core ideas, concepts, and practices of the C3 Inquiry Cycle as appropriate.
- ❒ Recommends and facilitates a mix of instructional approaches and best practices for a variety of learners, such as modeling, questioning strategies, checking for understanding, flexible grouping, pair-share, and scaffolding.
- ❒ Elicits students' prior knowledge and addresses common student conceptions/misconceptions.
- ❒ Supports students in making and evaluating evidence-based claims.
- ❒ Uses digital tools and media as appropriate to deepen student learning.
- ❒ Contains text features as appropriate to support student learning.
- ❒ Requires student involvement in and responsibility for their learning.

A longer lesson or unit should also:

- ❒ Demonstrate effective sequencing where the concepts and skills advance and deepen over time.
- ❒ Provide for various approaches to learning: relevant and transferable skills, application of literacy skills, student-directed inquiry, analysis, evaluation, and reflection.
- ❒ Use appropriate scaffolding, supporting student progress towards independent learning (may be more applicable across several units or the year).

Summary of Observations and Suggestions for Improvement:

Rating: 3 2 1 0

Rating Scale for Dimensions I, II, III, IV:

3: Meets most to all of the criteria in the dimension.
2: Meets many of the criteria in the dimension.
1: Meets some of the criteria in the dimension.
0: Does not meet the criteria in the dimension.

Unit Title: _____ **Reviewer ID:** _____

IV. Assessment

The lesson/unit provides sufficient guidance for interpreting student performance.

- ❒ Elicits evidence that a student can independently demonstrate that they can meet the targeted Washington State or other standard(s) identified in the lesson/unit.
- ❒ Includes aligned rubrics or assessment guidelines for interpreting student performance.
- ❒ Measures progress of fundamental understandings through embedded formative assessments that focus on learning target(s) and/or engage students in self-reflection.

A longer lesson or unit should also:

- ❒ Use varied modes of curriculum-embedded assessments that may include pre-, formative, summative, and self-assessment measures.

Summary of Observations and Suggestions for Improvement:

Rating: 3 2 1 0

Rating Scale for Dimensions I, II, III, IV:

3: Meets most to all of the criteria in the dimension.
2: Meets many of the criteria in the dimension.
1: Meets some of the criteria in the dimension.
0: Does not meet the criteria in the dimension.

Unit Title: _____ **Reviewer ID:** _____

Total Score:
Overall Rating:

Summary Comments

Overall Rating for the Lesson/Unit:

E: Exemplar - Aligned and meets most to all of the criteria in dimensions **(total 11-12)**

E/I: Exemplar *if* Improved – Aligned and needs some improvement in one or more dimensions **(total 8-10).**

R: Revision Needed – Aligned partially and needs significant revision in one or more dimensions **(total 3 – 7)**

N: Not Ready to Review – Not aligned and does not meet criteria **(total 0-2)**

version 2.3
updated 5/31/2016

CRITERIA for PROCURING and EVALUATING HIGH-QUALITY ASSESSMENTS

States have demonstrated their leadership and commitment to ensuring the success of all students by adopting college- and career-readiness standards. To realize the potential of these standards, states require assessments that match the depth, breadth, and rigor of the standards; accurately measure student progress toward college and career readiness; and provide valid data to inform teaching and learning.

Assessment of College and Career Readiness. States have taken different approaches to establishing college- and career-readiness standards and to putting in place high-quality aligned assessments. Many states have adopted the Common Core State Standards (CCSS); some have modified the CCSS to meet their state's context and needs; and others have developed standards independent of the CCSS. To provide assessments that are aligned to these standards, many states are working together through assessment consortia, while others are taking alternative paths for transition. This document is grounded in best practices for assessment development and in the research that defines college and career readiness for English Language Arts(ELA)/literacy and mathematics. Thus, regardless of each state's approach, this document is intended to be a useful resource for any state procuring and/or evaluating assessments aligned to their college- and career-readiness standards.

Assessment Criteria for States to Consider. This document provides criteria for states to consider as they develop procurements and evaluate options for high-quality state summative assessments aligned to college- and career-readiness standards. The criteria build on the states' high-quality summative assessment principles (CCSSO, 2013) which articulate their commitment to high-quality assessments aligned to college and career readiness. To assist states in operationalizing their commitment, this document pays particular attention to not only the criteria states could ask vendors to meet, but also to the evidence states could ask vendors to provide to demonstrate criteria have been – or will be – met. States will, of course, adapt these criteria to reflect their context, standards, and procurement regulations.

Contents of this Document. This document begins with an overview of the assessment criteria and continues with a chart containing detailed criteria and sample evidence. These criteria do not cover every area that a state would have to address in a procurement or evaluation process. Instead, they focus on the critical characteristics that should be met by high-quality assessments aligned to college- and career-readiness standards. A more comprehensive source for the development and validation of assessments is the *Standards for Educational and Psychological Testing* (AERA, APA, and NCME, 1999). The assessment criteria and evidence discussed herein were developed by referencing the *Standards for Educational and Psychological Testing* and several other key sources listed in the bibliography. Additional state-specific criteria at the end of the document highlight a few of the most important additional issues that states may wish to consider in a procurement or evaluation process.

Notes about Evidence and Terminology. This document is intended to support states in selecting assessments that meet a high bar for quality. Thus, the document suggests the evidence that states will need to review in order to make informed judgments on vendors' claims about the quality of their proposed assessments. Of course, vendors may propose assessments that are yet to be developed, assessments in development, and/or existing assessments. In designing procurement or evaluation procedures, states may therefore find it helpful to design the process for awarding "points" so as neither to reward existing (but poor quality) tests just because they have data available, nor to reward well-intentioned conceptual designs that are not executable. To support this goal, vendors should be asked to provide the most rigorous level of evidence they have available, consistent with the stage of assessment development they are in. The types of evidence that vendors should be expected to provide at different stages of development are described below:

- For assessments to be newly created, the most rigorous level of evidence will include the vendor's descriptions of their established and proven processes; data from similar assessments; proposed test blueprints and other specifications (e.g., test design documents, test specifications, item specifications, scoring specifications); exemplar test items, passages, and forms; proposed studies, reports, and technical documentation to be created during assessment development and operation; and the processes for responding to such data. In addition, the vendor's prior experience, expertise, and letters of recommendation should be included.
- For assessments that are currently in development, the most rigorous level of evidence will depend on the stage of assessment development. Evidence should include test blueprints and other specifications (e.g., test design documents, test specifications, item specifications, scoring specifications), and exemplar test items, passages, and forms. In addition, evidence should include as much of the data described below regarding pre-existing assessments as is available. Where such evidence is not available, vendors should provide descriptions of their established and proven processes; data from similar assessments, proposed studies, reports, and technical documentation to be created during assessment development and operation; and the process for responding to such data. In addition, the vendor's prior experience, expertise, and letters of recommendation should be included.
- For pre-existing assessments, the most rigorous level of evidence will include comprehensive validity evidence; test blueprints and other specifications (e.g., test design documents, test specifications, item specifications, scoring specifications); annual technical reports; results of studies on scaling, equating, and reporting; and exemplar test items, passages, and forms.

Additionally, regardless of the stage of test development, states may find it helpful to put in place best practice quality assurance and other processes so that states can monitor quality throughout development and administration, and periodically evaluate evidence to ensure criteria are being met.

Finally, a note about terminology. In this document, the term "assessments" generally refers to the entire suite of summative assessments a state would procure – that is, tests of ELA/literacy and mathematics in each grade assessed. In sections specifically about ELA/literacy or mathematics, however, the term refers to the set of summative assessments in that content area. The terms "assessment" and "test" are often used interchangeably when discussing a single grade level/content area. Throughout the document, the term "tasks" refers to extended-response, open-ended test items; "test items" refers to the stimuli used to elicit a response through, for example, multiple-choice or constructed-response items as well as tasks; and "forms" are systematic collections of test items and tasks that comprise the testing experience for a particular student in a grade/content area.

Overview of Assessment Criteria

A. Meet Overall Assessment Goals and Ensure Technical Quality

A.1 Indicating progress toward college and career readiness
A.2 Ensuring that assessments are valid for required and intended purposes
A.3 Ensuring that assessments are reliable
A.4 Ensuring that assessments are designed and implemented to yield valid and consistent test score interpretations within and across years
A.5 Providing accessibility to *all* students, including English learners and students with disabilities
A.6 Ensuring transparency of test design and expectations
A.7 Meeting all requirements for data privacy and ownership

B. Align to Standards – English Language Arts/Literacy

B.1 Assessing student reading and writing achievement in both ELA and literacy
B.2 Focusing on complexity of texts
B.3 Requiring students to read closely and use evidence from texts
B.4 Requiring a range of cognitive demand
B.5 Assessing writing
B.6 Emphasizing vocabulary and language skills
B.7 Assessing research and inquiry
B.8 Assessing speaking and listening
B.9 Ensuring high-quality items and a variety of item types

C. Align to Standards – Mathematics

C.1 Focusing strongly on the content most needed for success in later mathematics
C.2 Assessing a balance of concepts, procedures, and applications
C.3 Connecting practice to content
C.4 Requiring a range of cognitive demand
C.5 Ensuring high-quality items and a variety of item types

D. Yield Valuable Reports on Student Progress and Performance

D.1 Focusing on student achievement and progress to readiness
D.2 Providing timely data that inform instruction

E. Adhere to Best Practices in Test Administration

E.1 Maintaining necessary standardization and ensuring test security

F. State Specific Criteria (as desired)

Sample criteria might include

- Requiring involvement of the state's K-12 educators and institutions of higher education
- Procuring a system of aligned assessments, including diagnostic and interim assessments
- Ensuring interoperability of computer-administered items

Assessment Criteria and Evidence

A. Meet Overall Assessment Goals and Ensure Technical Quality*

Criteria	Evidence
A.1 **Indicating progress toward college and career readiness:** Scores and performance levels on assessments are mapped to determinations of college and career readiness at the high school level and for other grades to being on track to college and career readiness by the time of high school graduation.	• A description is provided of the process for developing performance level descriptors and setting performance standards (i.e., "cut scores"), including ○ Appropriate involvement of higher education and career/technical experts in determining the score at which there is a high probability that a student is college and career ready; ○ External evidence used to inform the setting of performance standards and a rationale for why certain forms of evidence are included and others are not (e.g., student performance on current state assessments, NAEP, TIMSS, PISA, ASVAB, ACT, SAT, results from Smarter Balanced and PARCC, relevant data on post-secondary performance, remediation, and workforce readiness); ○ Evidence and a rationale that the method(s) for including external benchmarks are valid for the intended purposes; and ○ Standard setting studies, the resulting performance level descriptors and performance standards, and the specific data on which they are based (when available). • A description is provided of the intended studies that will be conducted to evaluate the validity of performance standards over time.
A.2 **Ensuring that assessments are valid for required and intended purposes:** Assessments produce data, including student achievement data and student growth data required under Title I of the Elementary and Secondary Education Act (ESEA) and ESEA Flexibility, that can be used to validly inform the following: • School effectiveness and improvement; • Individual principal and teacher effectiveness for purposes of evaluation and identification of professional development and support needs; • Individual student gains and performance; and • Other purposes defined by the state.	• A well-articulated validity evaluation based on an interpretive argument (e.g., Kane, 2006) is provided that includes, at a minimum ○ Evidence of the validity of using results from the assessments for the three primary purposes, as well as any additional purposes required by the state (specify sources of data). ○ Evidence that scoring and reporting structures are consistent with structures of the state's standards (specify sources of data). ○ Evidence that total test and relevant sub-scores are related to external variables as expected (e.g., other measures of the construct). To the extent possible, include evidence that the items are "instructionally sensitive," that is, that item performance is more related to the quality of instruction than to out-of-school factors such as demographic variables. ○ Evidence that the assessments lead to the intended outcomes (i.e., meet the intended purposes) and minimize unintended negative consequences. Consequential evidence

* The term "technical quality" here refers to the qualities necessary to ensure that scoring and generalization inferences based on test scores are valid both within and across years. This document prioritizes certain aspects of technical quality, but as noted in the introduction, readers should also refer to other sources, primarily *The Standards for Educational and Psychological Testing*.

Criteria	Evidence
	should flow from a well-articulated theory of action about how the assessments are intended to work and be integrated with the larger accountability system. ○ The set of content standards against which the assessments are designed is provided. If these standards are the state's standards, evidence is provided that the content of the assessments reflects the standards, including the cognitive demand of the standards. If they are not the state's standards, evidence is provided of the extent of alignment with the state's standards. ○ Evidence is provided to ensure the content validity of test forms and the usefulness of score reports (e.g., test blueprints demonstrate the learning progressions reflected in the standards, and experts in the content and progression toward readiness are significantly involved in the development process).
A.3 **Ensuring that assessments are reliable:** Assessments minimize error that may distort interpretations of results, estimate the magnitude of error, and inform users of its magnitude.	• Evidence is provided of the reliability of assessment scores, based on the state's student population and reported subpopulations (specify sources of data). • Evidence is provided that the scores are reliable for the intended purposes for essentially all students, as indicated by the standard error of measurement across the score continuum (i.e., conditional standard error). • Evidence is provided of the precision of the assessments at cut scores, and consistency of student level classification (specify sources of data). • Evidence is provided of generalizability for all relevant sources, such as variability of groups, internal consistency of item responses, variability among schools, consistency from form to form of the test, and inter-rater consistency in scoring (specify sources of data).
A.4 **Ensuring that assessments are designed and implemented to yield valid and consistent test score interpretations within and across years:**	
• **Assessment forms** yield consistent score meanings over time, forms within year, student groups, and delivery mechanisms (e.g., paper, computer, including multiple computer platforms).	• A description is provided of the process used to ensure comparability of assessments and assessment results across groups and time. • Evidence is provided of valid and reliable linking procedures to ensure that the scores derived from the assessments are comparable within year across various test "forms" and across time. • Evidence is provided that the linking design and results are valid for test scores across the achievement continuum.
• **Score scales** used facilitate accurate and meaningful inferences about test performance.	• Evidence is provided that the procedures used to transform raw scores to scale scores is coherent with the test design and the intended claims, including the types of Item Response Theory (IRT) calibration and scaling methods (if used) and other methods for facilitating meaningful score interpretations over tests and time. • Evidence is provided that the assessments are designed and scaled to ensure the primary

Criteria	Evidence
	interpretations of the assessment can be fulfilled. For example, if the assessments are used as data sources for growth or value-added models for accountability purposes, evidence should be provided that the scaling and design features would support such uses, such as ensuring appropriate amounts of measurement information throughout the scale, as appropriate. • Evidence is provided, where a vertical or other score scale is used, that the scaling design and procedures lead to valid and reliable score interpretations over the full length of the scale proposed; and evidence is provided that the scale is able to maintain these properties over time (or a description of the proposed procedures is provided).
A.5 **Providing accessibility to *all* students, including English learners and students with disabilities:**	
• **Following the principles of universal design:** The assessments are developed in accordance with the principles of universal design and sound testing practice, so that the testing interface, whether paper- or technology-based, does not impede student performance.	• A description is provided of the item development process used to reduce construct irrelevance (e.g., eliminating unnecessary clutter in graphics, reducing construct-irrelevant reading load as much as possible), including ○ The *test item* development process to remove potential challenges due to factors such as disability, ethnicity, culture, geographic location, socioeconomic condition, or gender; and ○ *Test form* development specifications that ensure that assessments are clear and comprehensible for all students. • Evidence is provided, including exemplar tests (paper and pencil forms or screen shots) illustrating principles of universal design.
• **Offering appropriate accommodations and modifications:** Allowable accommodations and modifications that maintain the constructs being assessed are offered where feasible and appropriate, and consider the access needs (e.g., cognitive, processing, sensory, physical, language) of the vast majority of students.	• A description is provided of the accessibility features that will be available, consistent with state policy (e.g., magnification, audio representation of graphic elements, linguistic simplification, text-to-speech, speech-to-text, Braille). • A description is provided of access to translations and definitions, consistent with state policy. • A description is provided of the construct validity of the available accessibility features with a plan that ensures that the scores of students who have accommodations or modifications that do not maintain the construct being assessed are not combined with those of the bulk of students when computing or reporting scores.
• Assessments produce valid and reliable scores for **English learners**.	• Evidence is provided that test items and accessibility features permit English learners to demonstrate their knowledge and abilities and do not contain features that unnecessarily prevent them from accessing the content of the item. Evidence should address: presentation, response, setting, and timing and scheduling (specify sources of data).
• Assessments produce valid and reliable scores for **students with disabilities**.	• Evidence is provided that test items and accessibility features permit students with disabilities to demonstrate their knowledge and abilities and do not contain features that

Criteria	Evidence
	unnecessarily prevent them from accessing the content of the item. Evidence should address: presentation, response, setting, and timing and scheduling (specify sources of data).

A.6 **Ensuring transparency of test design and expectations:** Assessment design documents (e.g., item and test specifications) and sample test questions are made publicly available so that all stakeholders understand the purposes, expectations, and uses of the college- and career-ready assessments.	• Evidence is provided, including test blueprints, showing the range of state standards covered, reporting categories, and percentage of assessment items and score points by reporting category.
	• Evidence is provided, including a release plan, showing the extent to which a representative sample of items will be released on a regular basis (e.g., annually) across every grade level and content area.
	• Sample items with annotations and answer rationales are provided.
	• Scoring rubrics for constructed-response items with sample responses are provided for each level of the rubric.
	• Item development specifications are provided.
	• Additional information is provided to the state to demonstrate the overall quality of the assessment design, including
	  ○ Estimated testing time by grade level and content area;
	  ○ Number of forms available by grade level and content area;
	  ○ Plan for what percentage of items will be refreshed and how frequently;
	  ○ Specifications for the various levels of cognitive demand and how each is to be represented by grade level and content area; and
	  ○ For ELA/Literacy, data from text complexity analyses.

A.7 **Meeting all requirements for data privacy and ownership:** All assessments must meet federal and state requirements for student privacy, and all data is owned exclusively by the state.	• An assurance is provided of student privacy protection, reflecting compliance with all applicable federal and state laws and requirements. • An assurance is provided of state ownership of all data, reflecting knowledge of state laws and requirements. • An assurance is provided that the state will receive all underlying data, in a timely and useable fashion, so it can do further analysis as desired, including, for example, achievement, verification, forensic, and security analyses. • A description is provided for how data will be managed securely, including, for example, as data is transferred between vendors and the state.

B. Align to Standards – English Language Arts/Literacy

Criteria	Evidence
B.1 **Assessing student reading and writing achievement in both ELA and literacy:** The assessments are English language arts and literacy tests that are based on an aligned balance of high-quality literary and informational texts.	• Test blueprints and other specifications as well as exemplar literary and informational passages are provided for each grade level, demonstrating the expectations below are met. • Texts are balanced across literary and informational text types and across genres, with more informational than literary texts used as the assessments move up in the grade bands, as the state's standards require. *For example, for common core aligned assessments, goals include* ○ *In grades 3-8, approximately half of the texts are literature and half are informational;* ○ *In high school, because comprehension of complex informational texts is crucial for readiness, texts are approximately one-third literature and two-thirds informational; and* ○ *In all grades, informational texts are primarily expository rather than narrative in structure, and in grades 6-12, informational texts are approximately one-third each literary nonfiction, history/social studies, and science/technical.* • Texts and other stimuli (e.g., audio, visual, graphic) are previously published or of publishable quality. They are content-rich, exhibit exceptional craft and thought, and/or provide useful information. • History/social studies and science/technical texts, specifically, reflect the quality of writing that is produced by authorities in the particular academic discipline.
B.2 **Focusing on complexity of texts:** The assessments require appropriate levels of text complexity; they raise the bar for text complexity each year so students are ready for the demands of college- and career-level reading no later than the end of high school. Multiple forms of authentic, previously published texts are assessed, including written, audio, visual, and graphic, as technology and assessment constraints permit.	• Text complexity measurements, exemplar literary and informational passages for each grade level, and other evidence (e.g., data, tools, procedures) are provided to demonstrate the expectations below are met. • At each grade, reading texts have sufficient complexity, and the average complexity of texts increases grade-by-grade, meeting college- and career-ready levels by the end of high school. • A rationale and evidence are provided for how text complexity is quantitatively and qualitatively measured and used to place each text at the appropriate grade level. *For example, for common core aligned assessments, goals include* ○ *Texts are placed in a grade **band** using at least one research-based quantitative measure;* ○ *Texts are placed at a grade **level** using a qualitative analysis measure, reflecting the expert judgment of educators; and* ○ *Most of the texts are placed within the grade **band** indicated by the quantitative*

Criteria	Evidence
	analysis, with exceptions usually found in high school literary texts.
B.3 **Requiring students to read closely and use evidence from texts:** Reading assessments consist of test questions or tasks, as appropriate, that demand that students read carefully and deeply and use specific evidence from increasingly complex texts to obtain and defend correct responses.	• Test blueprints and other specifications as well as exemplar test items are provided for each grade level, demonstrating the expectations below are met. • All reading questions are text-dependent and o Arise from and require close reading and analysis of text; o Focus on the central ideas and important particulars of the text, rather than on superficial or peripheral concepts; and o Assess the depth and specific requirements delineated in the standards at each grade level (i.e., the concepts, topics, and texts specifically named in the grade-level standards). • Many reading questions require students to directly provide textual evidence in support of their responses. *For example, for common core aligned assessments, goals include* o *A majority of reading score points is devoted to questions that ask students to directly provide textual evidence in support of their responses (e.g., constructed-response and/or two-part evidence-based selected-response item formats).*
B.4 **Requiring a range of cognitive demand:** The assessments require all students to demonstrate a range of higher-order, analytical thinking skills in reading and writing based on the depth and complexity of college- and career-ready standards, allowing robust information to be gathered for students with varied levels of achievement.	• Test blueprints and other specifications are provided to demonstrate that the distribution of cognitive demand for each grade level and content area is sufficient to assess the depth and complexity of the state's standards, as evidenced by use of a generic taxonomy (e.g., Webb's Depth of Knowledge) or, preferably, classifications specific to the discipline and drawn from the requirements of the standards themselves and item response modes, such as o The complexity of the text on which an item is based; o The range of textual evidence an item requires (how many parts of text[s] students must locate and use to response to the item correctly); o The level of inference required; and o The mode of student response (e.g., selected-response, constructed-response). • A rationale is provided justifying the distribution of cognitive demand for each grade level and content area. • Exemplar test items for each grade level are provided, illustrating each level of cognitive demand, and accompanied by a description of the process used to determine an item's cognitive level.
B.5 **Assessing writing:** Assessments emphasize writing tasks that require students to engage in close reading and analysis of texts so that students can demonstrate college- and career-ready abilities.	• Test blueprints and other specifications as well as exemplar test items for each grade level are provided, demonstrating the expectations below are met. • Writing tasks reflect the types of writing that will prepare students for the work required in college and the workplace, balancing expository, persuasive/argument, and narrative writing, as state standards require. At higher grade levels, the balance shifts toward more exposition and argument.

Criteria	Evidence
	For example, for common core aligned assessments, goals include ○ *Taking all forms of the test together, writing tasks are approximately one-third each exposition, argument, and narrative (some tasks may represent blended structures), with the balance shifting toward more exposition and argument at the higher grade levels.* • Tasks (including narrative tasks) require students to confront text or other stimuli directly, to draw on textual evidence, and to support valid inferences from text or stimuli.
B.6 **Emphasizing vocabulary and language skills:** The assessments require students to demonstrate proficiency in the use of language, including vocabulary and conventions.	• Test blueprints and other specifications as well as exemplar test items for each grade level are provided, demonstrating the expectations below are met. • Vocabulary items reflect requirements for college and career readiness, including ○ Focusing on general academic (tier 2) words; ○ Asking students to use context to determine meaning; and ○ Assessing words that are important to the central ideas of the text. • Language is assessed within writing assessments as part of the scoring rubric, or it is assessed with test items that specifically address language skills. Language assessments reflect requirements for college and career readiness by ○ Mirroring real-world activities (e.g., actual editing or revision, actual writing); and ○ Focusing on common student errors and those conventions most important for readiness. • Assessments place sufficient emphasis on vocabulary and language skills (i.e., a significant percentage of the score points is devoted to these skills).
B.7 **Assessing research and inquiry:** The assessments require students to demonstrate research and inquiry skills, demonstrated by the ability to find, process, synthesize, organize, and use information from sources.	• Test blueprints and other specifications as well as exemplar test items for each grade level are provided, demonstrating the expectations below are met. • Test items assessing research and inquiry mirror real world activities and require students to analyze, synthesize, organize, and use information from sources. *For example, for common core aligned assessments, goals include* ○ *Research tasks require writing to sources, including analyzing, selecting, and organizing evidence from more than one source, and often from sources in diverse formats; and* ○ *When assessment constraints permit, real or simulated research tasks comprise a significant percentage of score points when all forms of the reading and writing test are considered together.*
B.8 **Assessing speaking and listening:** Over time, and as assessment advances allow, the assessments measure the speaking and listening communication skills students need for college and career readiness**.**	• Over time, and as assessment advances allow, the speaking and listening skills required for college and career readiness are assessed. *For example, for common core aligned assessments, test items assessing speaking* ○ *Assess students' ability to express well-supported ideas clearly and to probe others' ideas; and*

Criteria	Evidence
	○ *Include items that measure students' ability to marshal evidence from research and orally present findings in a performance task.* *For example, for common core aligned assessments, test items assessing listening* ○ *Are based on texts and other stimuli that meet the criteria for complexity, range, and quality outlined in criteria B.1 and B.2 above; and* ○ *Permit the evaluation of active listening skills (e.g., taking notes on main ideas, elaborating on remarks of others).*
B.9 **Ensuring high-quality items and a variety of item types:** High-quality items and a variety of types are strategically used to appropriately assess the standard(s).	• Specifications are provided to demonstrate that the distribution of item types for each grade level and content area is sufficient to strategically assess the depth and complexity of the standards being addressed. Item types may include, for example, selected-response, two-part evidence-based selected-response, short and extended constructed-response, technology-enhanced, and performance tasks. • To support claims of quality, the following are provided: ○ Exemplar items for each item type used in each grade band; ○ Rationales for the use of the specific item types; ○ Specifications showing the proportion of item types on a form; ○ For constructed response and performance tasks, a scoring plan (e.g., machine-scored, hand-scored, by whom, how trained), scoring rubrics, and sample student work to confirm the validity of the scoring process; and ○ A description of the process used for ensuring the technical quality, alignment to standards, and editorial accuracy of the items.

C. Align to Standards – Mathematics

Criteria	Evidence
C.1 **Focusing strongly on the content most needed for success in later mathematics:** The assessments help educators keep students on track to readiness by focusing strongly on the content most needed in each grade or course for later mathematics.	• Test blueprints and other specifications are provided, demonstrating that the vast majority of score points in each assessment focuses on the content that is most important for students to master in that grade band in order to reach college and career readiness. For each grade band, this content consists of ○ Elementary grades – number and operations; ○ Middle school – ratio, proportional relationships, pre-algebra, and algebra; and ○ High school – prerequisites for careers and a wide range of postsecondary studies, particularly algebra, functions, and modeling applications. *For example, for common core aligned assessments, goals include* ○ *In elementary grades, at least three-quarters of the points in each grade align exclusively to the major work of the grade;* ○ *In middle school grades, at least two-thirds of the points in each grade align exclusively*

Criteria	Evidence
	to the major work of the grade; and ○ *In high school, at least half of the points in each course align exclusively to prerequisites for careers and a wide range of postsecondary studies.* • The assessment design reflects the state's standards and reflects a coherent progression of mathematics content from grade to grade and course to course.
C.2 **Assessing a balance of concepts, procedures, and applications:** The assessments measure conceptual understanding, fluency and procedural skill, and application of mathematics, as set out in college- and career-ready standards.	• Test blueprints and other specifications as well as exemplar test items for each grade level are provided, demonstrating the expectations below are met. • The distribution of score points reflects a balance of mathematical concepts, procedures/fluency, and applications, as the state's standards require. *For example, for common core aligned assessments, at least one-quarter of the points come from each of the following categories:* ○ *Conceptual understanding problems in which students to respond to well-designed conceptual problems;* ○ *Procedural skill and fluency problems (e.g., purely procedural problems, some requiring use of efficient algorithms, and others inviting opportunistic strategies); and* ○ *Application problems (e.g., in elementary and middle grades, solving grade-appropriate word problems reflecting growing complexity across the grades; in high school, rich application problems requiring students to demonstrate college and career readiness).* • All students, whether high performing or low performing, are required to respond to items within the categories of conceptual understanding, procedural skill and fluency, and applications, so they have the opportunity to show what they know and can do.
C.3 **Connecting practice to content:** The assessments include brief questions and also longer questions that connect the most important mathematical content of the grade or course to mathematical practices, for example, modeling and making mathematical arguments.	• Test blueprints and other specifications as well as exemplar test items for each grade level are provided, demonstrating the expectations below are met. • Assessments for each grade and course meaningfully connect mathematical practices and processes with mathematical content (especially with the most important mathematical content at each grade), as required by the state's standards. • Explanatory materials (citing test blueprints and other specifications) describe the connection for each grade or course between content and mathematical practices and processes. *For example, for common core aligned assessments, goals include* ○ *Every test item that assesses mathematical practices is also aligned to one or more content standards (most often within the major work of the grade); and* ○ *Through the grades, test items reflect growing sophistication of mathematical practices with appropriate expectations at each grade level.*

Criteria	Evidence
C.4 **Requiring a range of cognitive demand:** The assessments require all students to demonstrate a range of higher-order, analytical thinking skills in mathematics based on the depth and complexity of college- and career-ready standards, allowing robust information to be gathered for students with varied levels of achievement. Assessments include questions, tasks, and prompts about the basic content of the grade or course as well as questions that reflect the complex challenge of college- and career-ready standards**.**	• Test blueprints and other specifications are provided to demonstrate that the distribution of cognitive demand for each grade level is sufficient to assess the depth and complexity of the state's standards, as evidenced by use a of generic taxonomy (e.g., Webb's Depth of Knowledge) or, preferably, classifications specific to the discipline and drawn from mathematical factors, such as o Mathematical topic coverage in the task (single topic vs. two topics vs. three topics vs. four or more topics); o Nature of reasoning (none, simple, moderate, complex); o Nature of computation (none, simple numeric, complex numeric or simple symbolic, complex symbolic); o Nature of application (none, routine word problem, non-routine or less well-posed word problem, fuller coverage of the modeling cycle); and o Cognitive actions (knowing or remembering, executing, understanding, investigating, or proving). • A rationale is provided justifying the distribution of cognitive demand for each grade level and content area. • Exemplar test items for each grade level are provided, illustrating each level of cognitive demand, and accompanied by a description of the process used to determine an item's cognitive level.
C.5 **Ensuring high-quality items and a variety of item types:** High-quality items and a variety of item types are strategically used to appropriately assess the standard(s).	• Specifications are provided to demonstrate that the distribution of item types for each grade level and content area is sufficient to strategically assess the depth and complexity of the standards being addressed. Item types may include selected-response, short and extended constructed-response, technology-enhanced, and multi-step problems. • To support claims of quality the following are provided: o The list and distribution of the types of work students will be asked to produce (e.g., facts, computation, diagrams, models, explanations); o Exemplar items for each item type used in each grade band; o Rationales for the use of the specific item types; o Specifications showing the proportion of item types on a form; o For constructed response items, a scoring plan (e.g., machine-scored, hand-scored, by whom, how trained), scoring rubrics, and sample student work to confirm the validity of the scoring process; and o A description of the process used for ensuring the technical quality, alignment to standards, and editorial accuracy of the items.

D. Yield Valuable Reports on Student Progress and Performance

Criteria	Evidence
D.1 **Focusing on student achievement and progress to readiness:** Score reports illustrate a student's progress on the continuum toward college and career readiness, grade by grade, and course by course. Reports stress the most important content, skills, and processes, and how the assessment focuses on them, to show whether or not students are on track to readiness.	• A list of reports is provided, and for each report, a sample that shows, at a minimum o Scores and sub-scores that will be reported with emphasis on the most important content, skills, and processes for each grade or course; o Explanations of results that are instructionally valuable and easily understood by essentially all audiences; o Results expressed in terms of performance standards (i.e., proficiency "cut scores"), not just scale scores or percentiles; and o Progress on the continuum toward college and career readiness, which can be expressed by whether a student has sufficiently mastered the current grade or course content and is therefore prepared for the next level. (Note: Not all reporting information need be numerical; for example, actual student work on a released item could be presented, along with the rubric for the item and a discussion of common errors.) • The reporting structure can be supported by the assessment design, as demonstrated by evidence, including data confirming that test blueprints include a sufficient number of items for each reporting category, so that scores and sub-scores lead to the intended interpretations and minimize the possibility of misinterpretation.
D.2 **Providing timely data that inform instruction:** Reports are instructionally valuable, easy to understand by all audiences, and delivered in time to provide useful, actionable data to students, parents, and teachers.	• A timeline and other evidence are provided to show when assessment results will be available for each report. • A description is provided of the process and technology that will be used to issue reports in as timely a manner as possible. • Evidence, including results of user testing, is provided to demonstrate the utility of the reports for each intended audience.

E. Adhere to Best Practices in Test Administration

Criteria	Evidence
E.1 **Maintaining necessary standardization and ensuring test security:** In order to ensure the validity, fairness, and integrity of state test results, the assessment systems maintain the security of the items and tests as well as the answer documents and related ancillary materials that result from test administrations.	• A comprehensive security plan is provided with auditable policies and procedures for test development, administration, score reporting, data management, and detection of irregularities consistent with NCES and CCSSO recommendations for, at a minimum o Training for all personnel – both test developers and administrators; o Secure management of assessments and assessment data, so that no individual gains access to unauthorized information; o Test administration and environment; and o Methods used to detect testing irregularities before, during, and after testing, and steps

Criteria	Evidence
	to address them. • A description is provided of how security safeguards have been tested and validated for computer-based tests and for paper-and-pencil tests, as relevant.

F. State Specific Criteria (as desired)

It is likely that states will supplement the above criteria with criteria specific to their needs. These might, for example, include

- **Requiring involvement of the state's K-12 educators, institutions of higher education, and career/technical experts** in the design, development, and/or scoring of the assessments;
- **Procuring a system of aligned assessments, including diagnostic and interim assessments** designed to target and improve instruction as well as measure progress and performance; and
- **Ensuring interoperability of computer-administered items** consistent in all ways with the specifications laid out in the *Assessment Interoperability Framework* (2012) developed by the Common Education Data Standards (CEDS) project, so that tests and items owned by the state can be easily ported from one technology platform to another.