

The Alignment between the *Iowa Assessments* and the Iowa Core

Form F

The Alignment between the Iowa Assessments and the Iowa Core

The *Iowa Assessments* have been designed, developed, researched and validated to support a variety of important educational purposes. These purposes include assessment of the Iowa Core, providing information about a student's performance relative to the state of Iowa's definition of proficiency, measuring a student's growth across time, providing information about a student's strengths and limitations, comparing students to the state and nation and informing changes in instruction and the evaluation of programs.

Content-oriented validity evidence is a very important component of many of these purposes. Evaluating the correspondence between the Iowa Core and test content provides information relative to the assessment of the Core as well as the state's definition of proficiency. In an annual assessment such as the Iowa Assessments, this evidence includes evaluating whether test content appropriately samples the domain set forward in Iowa Core, whether the cognitive demands of test items correspond to the level reflected in the Iowa Core, and whether the test avoids the inclusion of features irrelevant to the standard that is intended target of each test item.

The procedures used to develop and revise test materials are the foundation for an assessment's content validity. Meaningful evidence related to inferences based on content and performance standards guides the design and development of the content of any assessment. In addition, the Iowa Assessments are designed and implemented according to established professional standards, in order to ensure that the assessments are measures of what they claim to be, following guidelines in the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). In addition to the development process, test developers document the extent to which the content domain of a test represents the domain defined in the test specifications. Test developers also provide evidence of the extent to which the test items and scoring criteria yield scores that represent the defined domain. Detailed information about the development process for the Iowa Assessments can be found in the Iowa Research and Development Guide (2015).

This document provides a summary of three different types of alignment information for Form F of the Iowa Assessments. The first uses alignment criteria proposed by the Council of Chief State School Officers (CCSSO, 2014) 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 second provides a description of the item and test development activities used to help validate alignment and links to item-by-item content classifications for Form F. The third reports the results of an alignment study completed after Form F had been assembled.

Alignment Criteria for Evaluating High Quality Assessments

Specifically designed to address the multiple purposes of measuring growth, readiness and student achievement on core content, the Iowa Assessments focus on the highest-priority skills and knowledge at each grade, addressing foundational as well as complex skills. This approach focuses the evaluation on the highest-priority skills and knowledge at each grade. The Iowa Assessments define the prioritized skills and competencies at each grade level through the test specifications. Ongoing in-depth studies of educational standards and research literature on what knowledge and skills are most valuable for large-scale assessment has been a part of the University of Iowa's research agenda for the past 80 years. This research informs the test specifications prior to the development of each form. Given the priorities of measuring growth, proficiency, standards and reliability for a given assessment, research suggests that the results are most interpretable when the tests measure core knowledge and skills in depth rather than undertaking a surface-level wider range of topics. The identification of evidence-based core knowledge and skills on which college and career readiness and success rely define the specifications of the Iowa Assessments.

Traditional alignment methodologies, such as Webb's alignment or the Survey of Enacted Curriculum offer the advantage of having been studied extensively, but they only address a fairly narrow range of test-standards alignment issues. Another important theme is that, even for assessments that concentrate on measuring attainment of core knowledge and skills, it is a practical impossibility to assess every possible element in depth in a reasonable time frame. However, when each element belongs to a cohesive knowledge and skill domain, careful strategic **sampling** of the domain permits valid and reliable inferences about an examinee's level of learning. Careful domain sampling enables tests of reasonable length and time to render technically sound educational measurements.

Also, the test specifications of the Iowa Assessments were developed and validated using **empirical research of growth and college readiness**. The standards describing what students should know and be able to do at various score ranges are based on analysis of student responses across a nationally representative sample of students as well as the entire population of students from the state of Iowa. The progression across score ranges provides an empirical indicator of whether students are performing well enough in relation to those standards to be considered ready for postsecondary opportunities.

Because of the broadness of the assessment and the limitations of existing approaches, the document entitled *Criteria for Procuring and Evaluating High-Quality Assessments*, proposed by CCSSO (2014) was applied. These criteria are broad enough to address all purposes served by an assessment intended to measure growth, readiness and achievement on core content. A summary of these results is provided in Table 1 for reading and Table 2 for mathematics.

Table 1: CCSSO Alignment Criteria Applied to Iowa Assessments, Form F, Reading

Criteria*	Evidence
<p>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.*</p>	<ul style="list-style-type: none"> • See Table A.1. • See Research and Development Guide (2015) University of Iowa • See source documentation at: http://itp.education.uiowa.edu/ia/ContentAndCognitiveSpec.aspx • Test blueprints and other specifications as well as exemplar literary and informational passages are provided for each grade level. • 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 grades. • Texts are previously published or of publishable quality. They are content-rich, exhibit exceptional craft and thought, and/or provide useful information.
<p>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.</p>	<ul style="list-style-type: none"> • See Appendix C and Table C.1. • See Research and Development Guide (2015) University of Iowa • 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. • 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.
<p>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.</p>	<ul style="list-style-type: none"> • See Table A.1. • See Research and Development Guide (2015) University of Iowa • 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 rise from and require close reading and analysis of text; focus on the central ideas and important particulars of the text, rather than

	<p>on superficial or peripheral concepts; and 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).</p>
<p>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.</p>	<ul style="list-style-type: none"> • See Tables B.1 and B.2. • See source documentation at: http://itp.education.uiowa.edu/ia/ContentAndCognitiveSpec.aspx • See Research and Development Guide (2015) University of Iowa • Test blueprints and other specifications provide the distribution of cognitive demand for each grade level and content.
<p>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).</p>	<ul style="list-style-type: none"> • Specifications 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. • Exemplar items for each item type used in each grade are provided.
<p><i>* For the purpose of ESEA Title 1 – only the reading assessment criteria are presented here.</i></p>	

Table 2: CCSSO Alignment Criteria Applied to Iowa Assessments, Form F, Mathematics

Criteria	Evidence
<p>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.</p>	<ul style="list-style-type: none"> • See Tables A.2, A.3, A.4 • See source documentation at: http://itp.education.uiowa.edu/ia/ContentAndCognitiveSpec.aspx • 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 <ul style="list-style-type: none"> ○ Grades 3-5 – number and operations; ○ Grades 6-8 – ratios, proportional relationships, pre-algebra and algebra and ○ Grade 11 – prerequisites for careers and a wide range of postsecondary studies, particularly algebra, functions, and modeling applications.
<p>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.</p>	<ul style="list-style-type: none"> • The distribution of score points reflects a balance of mathematical concepts, procedures/fluency, and applications, at least one-quarter of the points come from each of the following categories: <ul style="list-style-type: none"> ○ Conceptual understanding problems ○ Procedural skill and fluency problems, and ○ Application problems • 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.

<p>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.</p>	<ul style="list-style-type: none"> • See source documentation at: http://itp.education.uiowa.edu/ia/ContentAndCognitiveSpec.aspx • 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. • Explanatory materials describe the connection for each grade or course between content and mathematical practices and processes.
<p>C.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.</p>	<ul style="list-style-type: none"> • See Tables B.1 and B.2. • See source documentation at: http://itp.education.uiowa.edu/ia/ContentAndCognitiveSpec.aspx • 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 Iowa Core.
<p>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).</p>	<ul style="list-style-type: none"> • 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 and Test Development Procedures to Help Ensure Alignment

In addition to the articulation of test specifications that define the content areas and cognitive processes to be measured by the Iowa Core, a well-defined set of procedures was used during item and test development stages to help ensure the alignment of the Iowa Assessments Form F to the Iowa Core. These procedures included:

- Training of item writers to create test materials (items and scoring rubrics) that measure critical aspects of the Iowa Core
- Alignment of individual items by educators to the Iowa Core during the item development and review processes. The process was an item-by-item evaluation of the content coverage and cognitive level of the items compared to the relevant domains of the Iowa Core
- Verification of alignment by a focus group of educators actively teaching the appropriate content areas at the appropriate grade level

The test specifications of the Iowa Assessments provide strong evidence of the categorical concurrence between the Iowa Assessments and the Iowa Core. All content categories identified for inclusion on the Iowa Assessments were consistent with domains identified within the Iowa Core. At the individual item level, each item on the Iowa Assessments was aligned to a standard within these domains from the Iowa Core by a minimum of seven educators during the item writing and item review processes. The results of the categorical concurrence between the Iowa Assessment items, the test specifications and the Iowa Core standards are available for use by Iowa educators. These results are posted on the Iowa Testing Programs website to promote the use of this information by Iowa educators.

The cognitive processes demanded by the items on the Iowa Assessments were consistent with the range of complexity found in the Iowa Core. To ensure this consistency, all items were reviewed by a minimum of seven educators during the item writing and item review processes for their cognitive demand to ensure that what students are expected to know and do is consistent between the two. The result was a full range of item complexity where each item in Form F was assigned one of three cognitive level descriptors. Table 6 in the appendix describes these three levels.

The range-of-knowledge criterion is used to judge whether the span of knowledge expected of students by a standard is the same as, or corresponds to, the span of knowledge that students need in order to answer correctly the test questions associated with that standard. The balance of representation indicates the degree to which the emphasis of the standards is reflected in the overall test specifications. To achieve the desirable range and balance, the Iowa Assessments have been designed to mirror the rigor of the Iowa Core while providing scores that are accurate around the two cut points of proficient/not proficient and advanced/proficient.

Post Test Assembly Validation of Alignment

After the assembly of Form F, an independent group of reviewers from the state of Iowa examined the strength of the relationship between the items on the Iowa Assessments and the Iowa Core, using a modified Webb-based protocol for studying categorical concurrence (Webb, 1999, 2002, 2007). The reviewers were selected based on their teaching experiences (both content and grade levels) as well as their experience implementing the Iowa Core in their classrooms. At the time of the study, all reviewers were currently teaching in Iowa schools. A minimum of four reviewers evaluated each item and a total of 22 reviewers participated in the study. Reviewers systematically compared each Iowa Assessments item in reading and mathematics to the standards of the Iowa Core, searching for the strongest match in terms of measuring a fundamental skill or concept embedded in one or more standards. For each Iowa Assessments item, reviewers assigned one of the following ratings to best describe the level of the match:

- *Strong alignment:* Reviewers found substantial or foundational overlap between the Iowa Assessments item and one or more standards in the Iowa Core. Both were found to measure the same central idea, fundamental skill, or core concept.
- *Alignment with multiple standards:* Depending upon the approach used by the student in responding to an Iowa Assessment item, reviewers could have identified two different standards could have been assessed.
- *No alignment:* Reviewers found no alignment between the Iowa Assessments item and any standards in the Iowa Core.

For this study, the items were pulled directly from Form F booklets of the Iowa Assessments. The Iowa Core was accessed through the Iowa Department of Education’s public website. As part of best practices for the alignment work, decision rules were developed and documented as necessary to ensure that procedures are transparent, systematic, and repeatable. This process ensures consistency over time and across analysts in their interpretations of items and standards text and in the application of the rating protocol.

The reviewers were instructed to align the items first to content standards from the Iowa Core. The number of mathematics and reading items aligned to one or multiple Iowa Core standards is provided in Table 3. All items on the Iowa Assessments were found to align with a minimum of one and frequently multiple standards found in the Iowa Core. The reviewers were also instructed to align the items to the appropriate grade level found in the Iowa Core standard. If they did not find an on-grade standard to align an item, they were to review the standards above and below that grade level. The number of items aligned above and below grade level included items that were both above and below grade level. Consistent with the purposes of the Iowa Assessments providing valid growth and readiness information, this finding was consistent with the overall design of the assessment to ensure the integrity of the vertical scale used to support valid growth and readiness information.

Overall, in reading and mathematics, reviewers found strong alignment between all items in the Iowa Assessments Form F and the Iowa Core.

Table 3: Post Test Assembly Alignment Results for Content

Content/ Grade	Total Number of Items	Number of Items Aligned to the Iowa Core by Reviewer (R)			
		R1	R2	R3	R4
Mathematics		R1	R2	R3	R4
3	50	50	49	50	46
4	55	55	51	54	52
5	60	60	54	59	59
6	65	63	60	65	64
7	70	68	67	70	68
8	75	75	71	74	75
11	40	40	38	40	39
Reading		R1	R2	R3	R4
3	41	40	41	41	41
4	42	41	42	42	41
5	43	43	43	43	41
6	44	43	44	44	43
7	45	44	45	45	44
8	46	46	46	44	45
11	40	40	40	39	39

Summary of Findings

At a general level, alignment between assessments and standards involves determining the degree to which standards of knowledge, skills and abilities that students should know and be able to with match, or align to, the measures selected to assess those standards. Underlying this match is the concept of validity. That is, the interpretations made about students' scores on tests can only be made in reference to the academic standards within the context of how those standards and the assessment align. With the Iowa Assessments as one such measure, the alignment criteria should be multifaceted and address the priorities of growth, readiness and student achievement. This document was developed to provide such evidence to the users of the Iowa Assessments.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Criteria for Procuring and Evaluating High-Quality Assessments; CCSSO (2014)

<http://www.ccsso.org/Documents/2014/CCSSO%20Criteria%20for%20High%20Quality%20Assessments%2003242014.pdf>

Webb, N. L. (1999). *Alignment of science and mathematics standards and assessments in four states* (Research Monograph No. 18). Washington, DC: Council of Chief State School Officers.

Webb, N. L. (2002). *An analysis of the alignment between mathematics standards and assessments for three states*. Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.

Webb, N. L. (2007). Issues related to judging the alignment of curriculum standards and assessments. *Applied Measurement in Education*, 20, 7–25

Appendix A – Content Specifications

Table A.1: Iowa Assessment – Reading Specifications -- All ESEA Grades

Iowa Core Domains in Literacy	Grades 3-8, 11 Numbers of Items						
	3	4	5	6	7	8	11
Reading Literature	20	20	21	16	16	17	8
Reading Informational Text	21	22	22	28	29	29	32
Key Ideas and Details	28	26	24	25	26	25	25
Craft and Structure	11	12	15	16	15	16	11
Integration of Knowledge and Ideas	2	4	4	3	4	5	4
Total	41	42	43	44	45	46	40

Table A.2: Iowa Assessment – Mathematics Specifications -- ESEA Grades 3-5

Iowa Core Domains in Mathematics	Grades 3-5 Numbers of Items		
	3	4	5
Number and Operations in Base Ten	6	8	11
Number and Operations—Fractions	2	2	2
Measurement and Data	22	22	23
Geometry	6	6	9
Operations and Algebraic Thinking	14	17	15
Total	50	55	60

Table A.3: Iowa Assessment – Mathematics Specifications -- ESEA Grades 6-8

Iowa Core Domains in Mathematics	Grades 6-8 Numbers of Items		
	6	7	8
Statistics & Probability	9	10	12
The Number System	19	17	16
Expressions & Equations	9	10	20
Geometry	14	15	17
Mathematical Practices	9	9	8
Ratios & Proportional Relationships	5	9	0
Functions	0	0	2
Total	65	70	75

Table A.4: Iowa Assessment – Mathematics Specifications -- ESEA Grade 11

Iowa Core Domains in Mathematics	Grade 11 Numbers of Items
Number and Quantity	12
Statistics & Probability	8
Algebra	10
Geometry	8
Functions	2
Total	40

Table A.5: Iowa Assessment – Science Specifications -- ESEA Grades 5, 8 and 11

Iowa Core Domains in Science	Grades 5, 8, and 11 Numbers of Items		
	5	8	11
Earth & Space Science	10	13	10
Life Science	15	15	19
Physical Science	12	15	19
Engineering, Technology, and Applications of Science	0	0	0
Total	37	43	48

Appendix B – Cognitive Specifications

Table B.1: Cognitive Descriptors – All ESEA Grades

Assessment	Essential competencies	Conceptual understanding	Extended reasoning
Mathematics	<ul style="list-style-type: none"> • Understand mathematical concepts or procedures 	<ul style="list-style-type: none"> • Make decisions of how to approach the problem • Specify and explain relationships between terms, properties, or operations • Perform multiple-step procedures 	<ul style="list-style-type: none"> • Use reasoning, use planning, draw conclusions, or cite evidence to solve a problem • Develop a strategy to connect and relate ideas to solve problems while using multiple-step procedures and a variety of skills
Reading	<ul style="list-style-type: none"> • Comprehend written text 	<ul style="list-style-type: none"> • Use more complex thought processes in interpreting and inferring from text • Determine important ideas 	<ul style="list-style-type: none"> • Use critical thinking in judging, evaluating, or analyzing text, or in integrating or synthesizing ideas within and beyond the text
Science	<ul style="list-style-type: none"> • Identify scientific information such as definitions, terminology, principles, concepts, and relationships • Recognize fundamental components of scientific investigations 	<ul style="list-style-type: none"> • Understand scientific concepts and apply them to explain phenomena • Analyze and interpret scientific information • Make simple inferences, predictions, and conclusions • Formulate hypotheses 	<ul style="list-style-type: none"> • Propose solutions to scientific problems • Make inferences, predictions, and conclusions • Evaluate appropriateness of scientific findings, conclusions, and experimental design • Integrate ideas from various sources

Table B.2: Iowa Assessment – Number of Items by Cognitive Classification -- All ESEA Grades

	Grades						
Reading	3	4	5	6	7	8	11
Essential Competencies	11	9	11	11	11	11	7
Conceptual Understandings	23	24	24	25	25	25	25
Extended Reasoning	7	9	8	8	9	10	8
Mathematics	3	4	5	6	7	8	11
Essential Competencies	6	6	6	6	6	6	5
Conceptual Understandings	39	43	47	52	59	62	29
Extended Reasoning	5	6	7	7	5	7	6

Appendix C – Text Complexity Process

Extensive processes are used to assess quantitative and qualitative assessment of text complexity for the Iowa Assessments. Three different dimensions of qualitative, quantitative, and reader/task considerations are used to describe the reading test. Table C.1 summarizes the type of information available to evaluate each dimension. All three dimensions were equally important in the assembly of Form F and were used to provide a range of text complexity and sophistication.

All text-based materials are reviewed by testing and content experts for four different aspects of the qualitative dimension including level of meaning or purpose, structure, language conventionality, and clarity. The quantitative dimensions are evaluated through a combination of text-based indices (for example, Lexiles and traditional readability indices) and national passage-based statistics that addresses the relative difficulty of these materials for the appropriate students. In addition, all passages are reviewed for accessibility, appropriateness of test complexity, and appropriateness of topics.

Table C.1: Evaluation of Text Complexity

Dimension	Reading	
Qualitative dimension	Levels of meaning or purpose	Includes a variety of literary and informational texts from simple meaning to multiple meanings
	Structure	Includes a variety of texts from simple to highly complex
	Language conventionality and clarity	Texts rely on a range of language conventionality and clarity from literal to figurative. Texts are balanced to represent this range within a given form of the HiSET.
	Knowledge demands	No assumptions about readers' life experiences
Quantitative dimension	Lexile scores for all text-based stimuli aligned to the college and career readiness ranges established by MetaMetrics Traditional readability indices for all text-based stimuli based on word length, frequency, and complexity Item-level and form-level difficulty indices collected from a nationally representative sample of students.	
Reader and task considerations	Student difficulty levels collected on nationally representative samples of students. Professional judgments from educators on the appropriateness of the passages and stimuli included in Form F.	

