



# **Review Criteria**

## **Interim Assessment**

### **English Language Arts**

#### **Grades 3-8**

v1.0

# Gateway 1

## Alignment, Fairness, & Accessibility

The assessment is designed to measure student knowledge and skills in alignment with the expectations of college- and career-ready standards.

The process for the development of items and test forms is clearly-articulated, intentional, and seeks to measure the depth and breadth of the standards.

The assessment is designed for fairness and accessibility; all students are provided opportunities to have their learning and achievement accurately assessed.

Gateway 1 Overview	Available Points
<b>Criterion 1.1: Test Development Alignment</b> Indicators 1.1.a-1.1.b  Assessment design specifications align to the expectations of college- and career-ready (CCR) standards.	8
<b>Criterion 1.2: Item and Form Alignment</b> Indicators 1.2.a-1.2.g  Text passages, assessment items, and resulting test forms align to the expectations of the ELA domains as outlined by college- and career-ready (CCR) standards.  <i>*Note: These indicators may be "N/C" if they are not claimed by the publisher to be present in the assessment.</i>	Points may vary based on indicators that are claimed by the publisher to be assessed.
<b>Criterion 1.3: Fairness and Accessibility</b> Indicators 1.3.a-1.3.c  The assessment is fair and accessible for all students in the intended test-taking population.	12

► Criterion 1.1: Test Development Alignment

Assessment design specifications align to the expectations of college- and career-ready (CCR) standards.

Indicators	Rating
<p>1.1.a Assessment design specifications provide clear expectations and detailed guidance to support the development of high-quality, CCR standards-aligned materials.</p> <ul style="list-style-type: none"> <li>Assessment rationale explains the design of the assessment, the benefits of the assessment, and a research foundation grounding the assessment process.</li> <li>Item development documentation is sufficiently robust to support the writing and review of items measuring CCR standards.</li> <li>Across all item types, assessment design specifications provide clear scoring information and/or rubrics to evaluate students' levels of understanding with respect to CCR standards being measured.</li> <li>Passage selection documentation details a process for the review and selection of texts based on the expectations of CCR standards.</li> <li>Item development documentation includes a description of processes used to ensure items are content-accurate and without technical or editorial flaws.</li> </ul>	0 2 4
<p>1.1.b Test blueprints and/or assessment design specifications reflect an appropriate distribution of content and related score points, item types, and cognitive demand within test events.</p> <ul style="list-style-type: none"> <li>The expected distribution of test content within and across ELA domains is defined within the test blueprint and/or assessment design specifications and reflects the emphasis established in CCR standards.</li> <li>The expected distribution of score points among ELA domains and subdomains is defined within the test blueprint and/or assessment design specifications and reflects the emphasis suggested in CCR standards.</li> <li>The expected type and range of item types is reflected in the test blueprint and/or assessment design specifications and are appropriate to address the expectations of CCR standards.</li> <li>The suggested ranges of cognitive demand are reflected in the test blueprint and/or assessment design specifications and are sufficient to measure the depth of CCR standards.</li> </ul>	0 2 4

Total  
Available  
Points

8

Meets: 8  
Partially Meets: 6  
Does Not Meet: <6

► Criterion 1.2: Item and Form Alignment

Text passages, assessment items, and resulting test forms align to the expectations of the ELA domains as outlined by college- and career-ready (CCR) standards.

Indicators	Rating
<p>1.2.a The text passages are of high quality and aligned with the expectations of CCR standards and aligned to assessment design specifications.</p> <ul style="list-style-type: none"> <li>• All texts or other stimuli included in the assessment are of publishable quality and are reflective of the requirements of CCR standards.</li> <li>• Text complexity is determined and documented using at least one research-based instrument; overall text complexity is determined through a combination of qualitative and quantitative analyses.</li> <li>• Texts are placed at the grade level indicated by the results of the text complexity process, with exceptions supported by explanation.</li> <li>• The distribution of texts and/or other stimuli are representative of the balance and range of text types required by CCR standards.</li> <li>• Texts provide adequate context for the design of multiple, meaningful items and are sufficiently developed to support logical inferences.</li> <li>• Selected texts align to the assessment design specifications.</li> </ul>	<p>0 2 4</p>
<p>1.2.b Test items are written to elicit evidence of learning relative to one or more of CCR standard/s and aligned to assessment design specifications.</p> <ul style="list-style-type: none"> <li>• Test items can be clearly identified as fully measuring one or more of CCR standard/s without formally measuring knowledge and skills that are not included within CCR standards.</li> <li>• Test items align to assessment design specifications.</li> <li>• Items are content-accurate and reflect no technical or editorial flaws.</li> </ul>	<p>0 2 4</p>
<p>*1.2.c The range of item types and cognitive demand among test events is sufficient to strategically assess the depth and complexity of CCR standards being addressed and is aligned to blueprints or assessment design specifications.</p> <ul style="list-style-type: none"> <li>• The item stimuli are constructed to reach the depth and complexity of CCR standards expressing multiple cognitive goals (e.g., determine and analyze; describe and explain; make connections among and distinctions between).</li> <li>• There is an appropriate distribution and/or range of cognitive demand exercised among test events submitted for review.</li> <li>• The range of item types and cognitive demand among test events align to blueprints or assessment design specifications.</li> <li>• If the assessment claims to measure writing, there is at least one extended constructed-response item to fulfill the expectations of CCR writing standards.</li> </ul>	<p>0 2 4</p> <p>N/C</p>
<p>*1.2.d The assessment is aligned to the reading expectations of CCR standards.</p> <ul style="list-style-type: none"> <li>• Reading items require students to use or provide textual evidence in support of responses requiring close reading and analysis (e.g., constructed-response and/or two-part evidence-based selected-response item formats).</li> <li>• Vocabulary items reflect the range of requirements for college and career readiness, including a focus on academic or Tier 2 words, the use of context to determine meaning, and an emphasis on words and phrases important to the central ideas of the text.</li> </ul>	<p>0 2 4</p> <p>N/C</p>



<p>*1.2.e The assessment is aligned to the writing, research, and language expectations of CCR standards.</p> <ul style="list-style-type: none"> <li>• Writing prompts and tasks represent the distribution reflected in CCR standards.</li> <li>• Writing tasks and/or prompts are text-based and require students to analyze, synthesize, organize, and write using evidence from a source or sources.</li> <li>• The majority of score points for items meeting the language standards are found in written responses or editing items that reflect authentic writing applications.</li> <li>• Items assessing conventions focus on common student errors that address conventions most critical for college and career readiness in real-world settings.</li> </ul>	<p>0   2   4</p> <p>N/C</p>
<p>*1.2.f The assessment is aligned to the speaking and listening expectations of the CCR standards.</p> <ul style="list-style-type: none"> <li>• Speaking items assess students' ability to draw on diverse content to prepare for, participate in, or orally present findings in a performance task.</li> <li>• Listening comprehension items assess students' ability to evaluate text and marshal information presented in diverse media forms.</li> </ul>	<p>0   2   4</p> <p>N/C</p>
<p>*1.2.g The assessment is aligned to the foundational skills expectations of reading CCR standards. (Grades 3-5 only)</p> <ul style="list-style-type: none"> <li>• Prompts and tasks represent the distribution of content reflected in the CCSS standards.</li> <li>• The assessment provides a variety of item elicitation formats appropriate for measuring CCR foundational skills.</li> </ul>	<p>0   2   4</p> <p>N/C</p>

*\*Note: These indicators may be "N/C" if they are not claimed by the publisher to be present in the assessment.*

<b>Total Available Points</b>	Points may vary based on indicators that are claimed by the publisher to be assessed.	<p>Meets: &gt;79% of applicable points</p> <p>Partially Meets: 50%-79% of applicable points</p> <p>Does Not Meet: &lt;50% of applicable points</p>
-------------------------------	---	--

► Criterion 1.3: Fairness and Accessibility

The assessment is fair and accessible for all students in the intended test-taking population.

Indicators	Rating
<p>1.3.a Items and test events are developed and reviewed using procedures that ensure fairness.</p> <ul style="list-style-type: none"> <li>Item development documentation/procedures clearly demonstrate adherence to the principles of universal design.</li> <li>Item rendering specifications clearly reflect the principles of universal design.</li> <li>Item review processes are designed to minimize construct-irrelevant variance.</li> <li>Items and passages go through a content bias/sensitivity review to make sure they are appropriate and fair for all relevant student groups.</li> <li>Procedures are in place to evaluate the technical quality and appropriateness of items and test events for student subgroups and students utilizing different accommodations.</li> </ul>	0 2 4
<p>1.3.b Appropriate accommodations and supports are in place to ensure the assessment is accessible to all students in the intended test-taking population, including special populations of students and English Learners.</p> <ul style="list-style-type: none"> <li>The test-taking population for which the assessment was/was not designed to support is clearly documented.</li> <li>The list of accommodations is aligned to the vendor's definition of the assessment's intended uses.</li> <li>The list of accommodations is sufficient to serve the needs of the full population of intended test takers.</li> <li>Evidence is available to support the validity and fairness of the intended interpretations and uses for those students who access the exam using the supported accommodations.</li> <li>Evidence is available that supports the quality and appropriateness of provided accommodations.</li> <li>The administration manual is clearly worded and supports teachers and other educational personnel in providing an appropriate testing experience for all students.</li> <li>Sample forms or released test items are available to stakeholders at each grade level.</li> </ul>	0 2 4
<p>1.3.c The range and types of technology provided within the assessment support the validity of assessment outcomes.</p> <ul style="list-style-type: none"> <li>Guidance is provided to support accessibility to the assessment system on a variety of platforms.</li> <li>Auditory supports present stimuli and items in a natural voice and at a cadence that can be adjusted to accommodate the learner.</li> <li>Overall visual design, including digital tools (e.g., dictionaries, thesauri, sticky notes, and highlighters) enhances the test-taking experience, does not distract or clutter the digital workspace, and can be easily navigated by students.</li> </ul>	0 2 4

Total  
Available  
Points

12

Meets: 10-12

Partially Meets: 6-9

Does Not Meet: <6

# Gateway 2

## Technical Quality

The assessment supports valid score interpretations related to student achievement and other reported information such as predicted student performance on future tests, areas of student strengths and need, and progress on student learning since previous test administrations.

Gateway 2 Overview	Available Points
<b>Criterion 2.1: Overall Achievement</b> Indicators 2.1.a-2.1.d  The interim assessment provides for valid inferences about a student's current overall achievement in the target content domain.	8
<b>Criterion 2.2: Predicted Student Performance</b> Indicators 2.2.a-2.2.d  The interim assessment provides valid information regarding predicted student performance on a state's summative assessment or other intended criterion measure(s).	Points may vary based on indicators that are claimed by the publisher to be assessed.
<b>Criterion 2.3: Sub-scores</b> Indicators 2.3.a-2.3.d  The interim assessment provides for valid inferences about a student's specific areas of strength and need (e.g., at the reportable category, content strand or objective level).	Points may vary based on indicators that are claimed by the publisher to be assessed.
<b>Criterion 2.4: Student Progress</b> Indicators 2.4.a-2.4.d  The interim assessment provides valid information regarding student progress in the content domain.	Points may vary based on indicators that are claimed by the publisher to be assessed.

► Criterion 2.1: Overall Achievement

The interim assessment provides for valid inferences about a student's current overall achievement in the target content domain.

Indicators	Rating
<p>2.1.a Item and form development procedures result in high-quality test events.</p> <ul style="list-style-type: none"> <li>Item development, review, and piloting procedures and materials are designed to ensure all newly developed items meet technical quality standards.</li> <li>Assessment design specifications and test development and review procedures ensure test events meet content and statistical quality criteria.</li> </ul>	0 1 2
<p>2.1.b Achievement scores are reliable.</p> <ul style="list-style-type: none"> <li>Item/test development and review procedures facilitate the reliability of test scores.</li> <li>Procedures for calculating and evaluating reliability are well-documented and appropriate.</li> <li>Obtained reliability indices and estimates of precision are at an appropriate level to support the use of results as intended.</li> </ul>	0 1 2
<p>2.1.c Achievement scores support intended interpretations of student performance.</p> <ul style="list-style-type: none"> <li>Evidence is provided to support the intended interpretations of student achievement.</li> <li>Equating/linking procedures supporting the comparability of achievement scores and score-based inferences across test events/administrations are described and reasonable.</li> <li>Item development specifications, task models, and scoring rubrics include enough detail to support consistency in the presentation, format, and degree of scaffolding observed in items and associated stimuli across test events.</li> <li>There is empirical evidence and an active research agenda supporting the validity of achievement scores as measures of the intended knowledge and skills.</li> </ul>	0 1 2
<p>2.1.d Achievement scores are appropriate for supporting their intended uses.</p> <ul style="list-style-type: none"> <li>The intended uses for the achievement scores are clearly and consistently articulated.</li> <li>There is sufficient theoretical and empirical evidence supporting the intended uses of achievement scores.</li> </ul>	0 1 2

Total Available Points

8

Meets: 7-8

Partially Meets: 5-6

Does Not Meet: <5

► Criterion 2.2: Predicted Student Performance

The interim assessment provides valid information regarding predicted student performance on a state's summative assessment or other intended criterion measure(s).

Indicators	Rating
<p>*2.2.a The design of the interim assessment supports its use in predicting performance on one or more external measures.</p> <ul style="list-style-type: none"> <li>• Sufficient information is provided to evaluate the degree to which the construct or content domain targeted by the interim assessment is similar to that assessed by the criterion measure(s).</li> <li>• The intended use of the interim assessment does not invalidate or contradict its appropriateness for predicting performance on the intended criterion measure(s).</li> <li>• If an interim assessment was designed to predict performance on specific assessments (e.g., ACT, SAT), evidence supporting that claim is provided.</li> </ul>	<p>0   1   2</p> <p>N/C</p>
<p>*2.2.b Predicted results are reliable.</p> <ul style="list-style-type: none"> <li>• Procedures used for calculating and evaluating the reliability of predicted scores/classifications are well documented and appropriate.</li> <li>• The reliability of the predicted result is calculated in a manner that is consistent with the inferences they were designed to support (e.g., CCR).</li> <li>• The predictions demonstrate sufficient reliability to support their intended uses.</li> </ul>	<p>0   1   2</p> <p>N/C</p>
<p>*2.2.c Predicted results (e.g., expected scaled scores, performance levels, passing status, etc.) reflect a student's likely performance on the state summative assessment or other intended criterion measure(s).</p> <ul style="list-style-type: none"> <li>• The data and procedures used to establish and evaluate the predictive relationship for a given test-taking sample are documented and reasonable.</li> <li>• The procedures used to support intended interpretations are clearly articulated.</li> <li>• Studies support the appropriateness of the predicted result as a measure of future performance.</li> </ul>	<p>0   1   2</p> <p>N/C</p>
<p>*2.2.d Predicted results are appropriate for supporting their intended uses.</p> <ul style="list-style-type: none"> <li>• The intended uses for the predicted results are clearly and consistently articulated.</li> <li>• There is sufficient theoretical and empirical evidence to support the appropriateness of the intended uses of predicted results.</li> </ul>	<p>0   1   2</p> <p>N/C</p>

*\*Note: Indicators and Criteria should be considered not claimed (N/C) if the assessment was not intentionally designed to be predictive.*

Total Available Points	Points may vary based on indicators that are claimed by the publisher to be assessed	Meets: Claim-Dependent  Partially Meets: Claim-Dependent  Does Not Meet: Claim-Dependent
------------------------	--	---

► Criterion 2.3: Sub-scores

The interim assessment provides for valid inferences about a student's specific areas of strength and need (e.g., at the reportable category, content strand or objective level).

Indicators	Rating
<p>*2.3.a Test events are designed to provide specific information about a student's areas of strength and need in the content domain.</p> <ul style="list-style-type: none"> <li>The assessment design supports the reporting of sub-scores at each level of granularity for which they are provided</li> <li>The assessment design supports interpretations of students' areas of strength and need in the content domain.</li> </ul>	<p>0 1 2</p> <p>N/C</p>
<p>*2.3.b Reported sub-scores are reliable.</p> <ul style="list-style-type: none"> <li>Estimates of reliability/precision are provided for all reported sub-scores.</li> <li>Procedures for calculating reliability indices and precision for the sub-score results are defensible and well documented.</li> <li>The calculated reliability and precision indices indicate adequate support for the intended interpretations and uses.</li> </ul>	<p>0 1 2</p> <p>N/C</p>
<p>*2.3.c Reported sub-scores support intended interpretations of student performance in defined sub-skill areas.</p> <ul style="list-style-type: none"> <li>Evidence is provided to support intended interpretations of all reported sub-scores.</li> <li>Empirical data suggest sub-scores represent distinct sub-domains and should be reported separately.</li> </ul>	<p>0 1 2</p> <p>N/C</p>
<p>*2.3.d Reported sub-scores are appropriate for supporting their intended uses.</p> <ul style="list-style-type: none"> <li>The intended uses for the sub-scores are clearly and consistently articulated.</li> <li>There is sufficient theoretical and empirical evidence supporting the intended uses for the sub-scores.</li> </ul>	<p>0 1 2</p> <p>N/C</p>

*\*Note: Indicators and Criteria should be considered not claimed (N/C) if the assessment was not intentionally designed with the use of subscores.*

Total Available Points	Points may vary based on indicators that are claimed by the publisher to be assessed	<p>Meets: Claim-Dependent</p> <p>Partially Meets: Claim-Dependent</p> <p>Does Not Meet: Claim-Dependent</p>
------------------------	--	---



► Criterion 2.4: Student Progress

The interim assessment provides valid information regarding student progress in the content domain.

Indicators	Rating
<p>*2.4.a The interim assessment is designed to support measures of growth.</p> <ul style="list-style-type: none"> <li>• Test design and content specifications (within and across grades) support the use of assessment results as a means of evaluating growth in the manner specified by the vendor.</li> <li>• The technical characteristics of the test and reportable scale support the reported growth measure.</li> </ul>	<p>0 1 2</p> <p>N/C</p>
<p>*2.4.b Student growth scores are reliable.</p> <ul style="list-style-type: none"> <li>• Procedures for estimating standard errors around the growth estimates are appropriate and well documented.</li> <li>• The reliability of the growth scores have been evaluated for students at different places along the ability scale.</li> <li>• The calculated reliability and precision indices indicate adequate support for the intended uses of the reported growth scores.</li> </ul>	<p>0 1 2</p> <p>N/C</p>
<p>*2.4.c Student growth scores support the intended interpretations.</p> <ul style="list-style-type: none"> <li>• The procedures and measures for calculating student growth are clearly documented and appropriate.</li> <li>• If significant modifications are made to the interim assessment that might break the trend line (i.e., test design changes, rescaling, and shifts in performance standards), empirical evidence is provided to support the intended interpretations and uses of growth scores.</li> <li>• Empirical evidence confirms that growth scores provide for valid intended inferences about student learning in the content domain.</li> </ul>	<p>0 1 2</p> <p>N/C</p>
<p>*2.4.d Student growth scores are appropriate for supporting the intended uses.</p> <ul style="list-style-type: none"> <li>• The intended uses for the growth scores are clearly and consistently articulated.</li> <li>• There is sufficient theoretical and empirical evidence supporting the intended uses for the growth scores.</li> </ul>	<p>0 1 2</p> <p>N/C</p>

*\*Note: Indicators and Criteria should be considered not claimed (N/C) if the assessment was not intentionally designed to support student progress.*

Total Available Points	Points may vary based on indicators that are claimed by the publisher to be assessed.	Meets: Claim-Dependent  Partially Meets: Claim-Dependent  Does Not Meet: Claim-Dependent
------------------------	---	---

# Gateway 3

## Score Reports and Interpretive Guides

Score reports and their resources provide sound information for stakeholders and support to assure assessment data is interpreted correctly and appropriately for use in multiple contexts.

Gateway 3 Overview	Available Points
<b>Criterion 3.1: Overall Achievement</b> Indicators 3.1.a-3.1.c  Score reports and other resources (e.g., user manuals, interpretive guides, instructional or curricular resources) are appropriate for facilitating the intended interpretations and uses of overall achievement results.	<b>10</b>
<b>Criterion 3.2: Predicted Student Performance</b> Indicators 3.2.a-3.2.c  Score reports and other resources (e.g., user manuals, interpretive guides, instructional or curricular resources) are appropriate for facilitating the intended interpretations and uses of predicted student performance.	Points may vary based on indicators that are claimed by the publisher to be assessed.
<b>Criterion 3.3: Sub-scores</b> Indicators 3.3.a-3.3.c  Score reports and other resources (e.g., user manuals, interpretive guides, instructional or curricular resources) are appropriate for facilitating the intended interpretations and uses of sub-scores.	Points may vary based on indicators that are claimed by the publisher to be assessed.
<b>Criterion 3.4: Student Progress</b> Indicators 3.4.a-3.4.c  Score reports and other resources (e.g., user manuals, interpretive guides, instructional or curricular resources) are appropriate for facilitating the intended interpretations and uses of student growth or progress results.	Points may vary based on indicators that are claimed by the publisher to be assessed.

► Criterion 3.1: Overall Achievement

Score reports and other resources (e.g., user manuals, interpretive guides, instructional or curricular resources) are appropriate for facilitating the intended interpretations and uses of overall achievement results.

Indicators	Rating
<p>3.1.a The design of the score reports and supporting materials (e.g., user manuals and interpretive guides) and the types of information provided are consistent with the intended interpretations and uses for specific users (e.g., educators, parents, students, or administrators).</p> <ul style="list-style-type: none"> <li>Score reports effectively represent the intended interpretations and uses of overall achievement results.</li> <li>The type and grain size of the information reported is appropriate for effectively serving the intended interpretations and uses.</li> <li>Evidence shows that there was attention to the audience and specific users in the design process, including user-specific versions of reports when needed.</li> <li>Evidence (e.g., studies, focus groups) is provided that users are able to effectively interpret and use reports in the manner intended.</li> <li>The documentation should include warnings of potential or common misuses of the results that may result in negative, unintended consequences.</li> <li>Reports identify and flag students for whom the integrity of the test interpretations may be compromised (e.g., student clicks through rapidly). <ul style="list-style-type: none"> <li>The conditions which bring about a flag are articulated on reports and/or in interpretive guides.</li> </ul> </li> </ul>	0 2 4
<p>3.1.b Score reports include information about the degree of error associated with the achievement score.</p> <ul style="list-style-type: none"> <li>For example, confidence intervals, error bands, or probability statements are provided to represent potential score variability.</li> <li>Supports (e.g., illustrative examples, informational text) are provided to facilitate accurate interpretations of error estimates and clarify the practical implications of error on score use.</li> </ul>	0 1 2
<p>3.1.c Sufficient and appropriate guidance (e.g., instructional or curricular supports) is provided to support the intended interpretations and uses, when needed.</p> <ul style="list-style-type: none"> <li>Guidance is aligned to the use.</li> <li>Any guidance provided has a basis in research and/or was created in consultation with educators experienced in using educational data.</li> <li>Guidance is provided to support appropriate use for students scoring at the full range of performance outcomes.</li> </ul>	0 2 4

Total Available Points

10

Meets: 8-10

Partially Meets: 5-7

Does Not Meet: <5

► Criterion 3.2: Predicted Student Performance

Score reports and other resources (e.g., user manuals, interpretive guides, instructional or curricular resources) are appropriate for facilitating the intended interpretations and uses of predicted student performance.

Indicators	Rating
<p><b>*3.2.a</b> The design of the score reports and supporting materials (e.g., user manuals and interpretive guides) and the types of information provided are consistent with the intended interpretations and uses for specific users (e.g., educators, parents, students, or administrators).</p> <ul style="list-style-type: none"> <li>Score reports effectively represent the intended interpretations and uses of overall achievement results.</li> <li>The type and grain size of the information reported is appropriate for effectively serving the intended interpretations and uses.</li> <li>Evidence shows that there was attention to the audience and specific users in the design process, including user-specific versions of reports when needed.</li> <li>Evidence (e.g., studies, focus groups) is provided that users are able to effectively interpret and use reports in the manner intended.</li> <li>The documentation should include warnings of potential or common misuses of the results that may result in negative, unintended consequences.</li> <li>Reports identify and flag students for whom the integrity of the test interpretations may be compromised (e.g., student clicks through rapidly). <ul style="list-style-type: none"> <li>The conditions which bring about a flag are articulated on reports and/or in interpretive guides.</li> </ul> </li> </ul>	<p>0 2 4</p> <p>N/C</p>
<p><b>*3.2.b</b> Score reports include information about the degree of error associated with the predicted performance score.</p> <ul style="list-style-type: none"> <li>For example, confidence intervals, error bands, or probability statements are provided to represent potential score variability.</li> <li>Supports (e.g., illustrative examples, informational text) are provided to facilitate accurate interpretations of error estimates and clarify the practical implications of error on score use.</li> </ul>	<p>0 1 2</p> <p>N/C</p>
<p><b>*3.2.c</b> Sufficient and appropriate guidance (e.g., instructional or curricular supports) is provided to support the intended interpretations and uses, when needed.</p> <ul style="list-style-type: none"> <li>Guidance is aligned to the use.</li> <li>Any guidance provided has a basis in research and/or was created in consultation with educators experienced in using educational data.</li> <li>Guidance is provided to support appropriate use for students scoring at the full range of performance outcomes.</li> </ul>	<p>0 2 4</p> <p>N/C</p>

*\*Note: These claims should match claims made and evaluated in Gateway 2. Indicators and Criteria should be considered not claimed (N/C) if the assessment was not intentionally designed to be predictive.*

Total Available Points	Points may vary based on indicators that are claimed by the publisher to be assessed.	Meets: Claim-Dependent  Partially Meets: Claim-Dependent  Does Not Meet: Claim-Dependent
------------------------	---	---

### ► Criterion 3.3: Sub-scores

Score reports and other resources (e.g., user manuals, interpretive guides, instructional or curricular resources) are appropriate for facilitating the intended interpretations and uses of sub-scores.

*Note: Some evidence statements in 3.3.a and 3.3.c may not be applicable due to previous review in Criterion 3.1.*

Indicators	Rating
<p><b>*3.3.a</b> The design of the score reports and supporting materials (e.g., user manuals and interpretive guides) and the types of information provided are consistent with the intended interpretations and uses for specific users (e.g., educators, parents, students, or administrators).</p> <ul style="list-style-type: none"> <li>Score reports effectively represent the intended interpretations and uses of overall achievement results.</li> <li>The type and grain size of the information reported is appropriate for effectively serving the intended interpretations and uses.</li> <li>Evidence shows that there was attention to the audience and specific users in the design process, including user-specific versions of reports when needed.</li> <li>Evidence (e.g., studies, focus groups) is provided that users are able to effectively interpret and use reports in the manner intended.</li> <li>The documentation should include warnings of potential or common misuses of the results that may result in negative, unintended consequences.</li> <li>Reports identify and flag students for whom the integrity of the test interpretations may be compromised (e.g., student clicks through rapidly). <ul style="list-style-type: none"> <li>The conditions which bring about a flag are articulated on reports and/or in interpretive guides.</li> </ul> </li> </ul>	<p>0 2 4</p> <p>N/C</p>
<p><b>*3.3.b</b> Score reports include information about the degree of error associated with sub-scores.</p> <ul style="list-style-type: none"> <li>For example, confidence intervals, error bands, or probability statements are provided to represent potential score variability.</li> <li>Supports (e.g., illustrative examples, informational text) are provided to facilitate accurate interpretations of error estimates and clarify the practical implications of error on score use.</li> </ul>	<p>0 1 2</p> <p>N/C</p>
<p><b>*3.3.c</b> Sufficient and appropriate guidance (e.g., instructional or curricular supports) is provided to support the intended interpretations and uses, when needed.</p> <ul style="list-style-type: none"> <li>Guidance is aligned to the use.</li> <li>Any guidance provided has a basis in research and/or was created in consultation with educators experienced in using educational data.</li> <li>Guidance is provided to support appropriate use for students scoring at the full range of performance outcomes.</li> </ul>	<p>0 2 4</p> <p>N/C</p>

*\*Note: These claims should match claims made and evaluated in Gateway 2. Indicators and Criteria should be considered not claimed (N/C) if the assessment was not intentionally designed with the use of subscores.*

Total Available Points	Points may vary based on indicators that are claimed by the publisher to be assessed.	Meets: Claim-Dependent  Partially Meets: Claim-Dependent  Does Not Meet: Claim-Dependent
------------------------	---	---

## ► Criterion 3.4: Student Progress

Score reports and other resources (e.g., user manuals, interpretive guides, instructional or curricular resources) are appropriate for facilitating the intended interpretations and uses of student growth or progress results.

Indicators	Rating
<p><b>*3.4.a</b> The design of the score reports and supporting materials (e.g., user manuals and interpretive guides) and the types of information provided are consistent with the intended interpretations and uses for specific users (e.g., educators, parents, students, or administrators).</p> <ul style="list-style-type: none"> <li>Score reports effectively represent the intended interpretations and uses of overall achievement results.</li> <li>The type and grain size of the information reported is appropriate for effectively serving the intended interpretations and uses.</li> <li>Evidence shows that there was attention to the audience and specific users in the design process, including user-specific versions of reports when needed.</li> <li>Evidence (e.g., studies, focus groups) is provided that users are able to effectively interpret and use reports in the manner intended.</li> <li>The documentation should include warnings of potential or common misuses of the results that may result in negative, unintended consequences.</li> <li>Reports identify and flag students for whom the integrity of the test interpretations may be compromised (e.g., student clicks through rapidly). <ul style="list-style-type: none"> <li>The conditions which bring about a flag are articulated on reports and/or in interpretive guides.</li> </ul> </li> </ul>	<p>0 2 4</p> <p>N/C</p>
<p><b>*3.4.b</b> Score reports include information about the degree of error associated with the student progress score.</p> <ul style="list-style-type: none"> <li>For example, confidence intervals, error bands, or probability statements are provided to represent potential score variability.</li> <li>Supports (e.g., illustrative examples, informational text) are provided to facilitate accurate interpretations of error estimates and clarify the practical implications of error on score use.</li> </ul>	<p>0 1 2</p> <p>N/C</p>
<p><b>*3.4.c</b> Sufficient and appropriate guidance (e.g., instructional or curricular supports) is provided to support the intended interpretations and uses, when needed.</p> <ul style="list-style-type: none"> <li>Guidance is aligned to the use.</li> <li>Any guidance provided has a basis in research and/or was created in consultation with educators experienced in using educational data.</li> <li>Guidance is provided to support appropriate use for students scoring at the full range of performance outcomes.</li> </ul>	<p>0 2 4</p> <p>N/C</p>

*\*Note: These claims should match claims made and evaluated in Gateway 2. Indicators and Criteria should be considered not claimed (N/C) if the assessment was not intentionally designed to support student progress.*

Total Available Points	Points may vary based on indicators that are claimed by the publisher to be assessed.	Meets: Claim-Dependent  Partially Meets: Claim-Dependent  Does Not Meet: Claim-Dependent
------------------------	---	---