

# **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](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](http://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.**

# Directions for Alignment Criterion 1

Reading – Range of Texts

## Alignment Criterion 1: Texts reflect the distribution of text types and genres required by the reading Standards.

### 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 acknowledgments
- 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
- Test blueprints and other explanatory material focused on test design, including sample score reports if available

Core State Standards. The more points the materials receive on the Alignment Criterion, the better the materials are aligned.

Evaluators should provide examples of evidence in support of the numerical rating for each metric, including evidence of any specific gaps in the materials.

### Rating this Criterion

Rate the assessments for each of the metrics for these Alignment Criteria as Meets (2 points), Partially Meets (1 point), or Does Not Meet (0 points). Then rate the Alignment Criteria as a group as Meets or Does Not Meet, based on the minimum number of points required. For the Alignment Criteria for reading, materials must earn at least 16 of 20 points to align to the Shifts and major features of the Common

# Alignment Criterion 1

## Reading – Range of Texts

### Metric

---

#### AC Metric 1A:

In 100% of the grades, the texts on reading assessments or in an item bank approximate the distributions of literary and informational texts as required by the Standards:

- In grades 3 – 8, there is a distribution of approximately 50%/50% literary and informational texts.
- In grades 9 – 12, there is a distribution of approximately 33% literary and 66% informational texts.

If the above metric is met, assign the materials 2 points. If 2 of the 3 grades within a grade band approximate the above distributions, assign 1 point.

### Procedure for Evaluation

---

At all grades, the proportions of literary vs. informational text should reflect the emphases in the Standards.

For each grade, examine the metadata accompanying either the texts on the test form(s) or blueprints or a random sample of at least 12 texts per grade.

List the texts and write “literary” or “informational” next to the title of each text. In accordance with the Standards, classify literary nonfiction texts as informational.

Calculate the percentages of literary vs. informational texts for each grade.

### Evidence

---

### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

# Alignment Criterion 1

## Reading – Range of Texts

### Metric

---

**AC Metric 1B:**

At least 90% of the literary and informational texts represent the genres and text characteristics that are specifically required by the reading Standards at each grade level.

If the above metric is met, assign the materials 2 points. If 60% to 90% of the texts meet the specific requirements of the Standards at each grade, assign 1 point.

### Procedure for Evaluation

---

At all grades, text types should match the Standards (e.g., specific genres and subgenres of fiction and nonfiction, foundational or seminal documents).

For each grade, examine the metadata accompanying the same texts as those used to evaluate the metrics in Non-Negotiable 1 above.

Write the genre or type next to each text on the list (e.g., “story,” “poem,” “literary nonfiction,” “science/technical,” “history/social studies”).

Compare the text characteristics to those required by the Standards at each grade and identify any texts that do not match the characteristics for that grade.

Calculate percentages.

### Evidence

---

### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

# Alignment Criterion 1

## Reading – Range of Texts

### Metric

---

#### AC Metric 1C:

Informational texts, specifically, meet the requirements of the Standards. At all grades, more than half of the informational texts utilize expository, rather than narrative, structures. In grades 6 – 12, the informational texts are balanced among history/social studies texts, science/technical texts, and literary nonfiction.

If the above metric is met, assign the materials 2 points. If one fourth to one half of the informational texts use expository rather than narrative structures and in grades 6 – 12 the informational texts include some history/ social studies, some science/technical, and some literary nonfiction, assign 1 point.

### Procedure for Evaluation

---

The ability to understand complex informational texts with expository structures is important for college and career readiness, as is the ability to understand complex informational texts within a variety of disciplines.

For each grade, note the primary structures in the informational texts in the list of texts used to evaluate the metrics in Non-Negotiable 1 above.

For grades 6 – 12, note the subject matter for the informational texts used to evaluate the metrics in Non-Negotiable 1 above.

Calculate whether more than half of the informational texts primarily use expository structures and in grades 6 – 12 whether there is a balance among history, science, and literary nonfiction.

### Evidence

---

### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

## Directions for Alignment Criterion 2

Reading – Assessing Vocabulary

**Alignment Criterion 2: Because of the importance of vocabulary acquisition and use to college and career readiness, vocabulary questions comprise a significant part of ELA/literacy assessments, assess tier 2 words in context, and focus on central ideas in the text.**

### 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 acknowledgments
- 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
- Test blueprints and other explanatory material focused on test design, including sample score reports if available

Core State Standards. The more points the materials receive on the Alignment Criterion, the better the materials are aligned.

Evaluators should provide examples of evidence in support of the numerical rating for each metric, including evidence of any specific gaps in the materials.

### Rating this Criterion

Rate the assessments for each of the metrics for these Alignment Criteria as Meets (2 points), Partially Meets (1 point), or Does Not Meet (0 points). Then rate the Alignment Criteria as a group as Meets or Does Not Meet, based on the minimum number of points required. For the Alignment Criteria for reading, materials must earn at least 16 of 20 points to align to the Shifts and major features of the Common

# Alignment Criterion 2

## Reading – Assessing Vocabulary

### Metric

---

**AC Metric 2A:**

At least 66% of vocabulary items emphasize the academic language that is crucial for readiness, and at least 90% require use of context to determine meaning.

If the above metric is met, assign the materials 2 points. If 40% to 66% of vocabulary items emphasize academic language, assign 1 point.

### Procedure for Evaluation

---

Most of the vocabulary items on assessments and in an item bank should assess academic vocabulary (tier 2) words or phrases in context. The remaining vocabulary items should assess other kinds of words named in the Standards (e.g., figurative and domain-specific language).

For each grade, examine either the vocabulary questions on the test form(s) or a representative sample (at least 15 vocabulary questions per grade) in the item bank.

List the sequence numbers of the questions that do not assess academic language (tier 2) words or phrases in context.

Calculate percentages.

### Evidence

---

### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

# Alignment Criterion 2

## Reading – Assessing Vocabulary

### Metric

---

**AC Metric 2B:**

At least 90% of vocabulary items assess words or phrases that are important to central ideas of the text.

If the above metric is met, assign the materials 2 points. If 50% to 90% of vocabulary items assess words or phrases important to central ideas, assign 1 point.

### Procedure for Evaluation

---

Vocabulary items on assessments and in an item bank should target words and phrases that are significant to the meaning of the text, not just unusual or interesting turns of phrase. The tested words or phrases should help students gain an understanding of the central ideas of a text, giving students a significant “payoff” when they determine the meaning.

For each grade, examine the vocabulary test questions assembled for Alignment Criterion 2A above.

List the sequence numbers of the questions that do not assess words that are important to the central ideas of the text.

Calculate percentages.

### Evidence

---

### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

# Alignment Criterion 2

## Reading – Assessing Vocabulary

### Metric

---

**AC Metric 2C:**

Vocabulary questions comprise a sufficient part of ELA/literacy assessments—at least 8 score points per test (which is a generally accepted minimum for a reporting category).

If the above metric is met, assign the materials 2 points. If 5 – 7 score points are given to vocabulary questions, assign 1 point.

### Procedure for Evaluation

---

At each grade, each assessment should include a sufficient number of points for vocabulary so that vocabulary could be a reporting category. Providing a reporting category for vocabulary is desirable but is not required.

For each grade, examine either the test blueprints or other test specifications

Determine the number of score points devoted to vocabulary per grade.

### Evidence

---

### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

# Directions for Alignment Criterion 3

Reading – Aligned Use of Item Types

**Alignment Criterion 3: A variety of item types is used to appropriately and strategically assess the Standards.**

## 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
- 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
- Test blueprints and other explanatory material focused on test design, including sample score reports if available

Core State Standards. The more points the materials receive on the Alignment Criterion, the better the materials are aligned.

Evaluators should provide examples of evidence in support of the numerical rating for each metric, including evidence of any specific gaps in the materials.

## Rating this Criterion

Rate the assessments for each of the metrics for these Alignment Criteria as Meets (2 points), Partially Meets (1 point), or Does Not Meet (0 points). Then rate the Alignment Criteria as a group as Meets or Does Not Meet, based on the minimum number of points required. For the Alignment Criteria for reading, materials must earn at least 16 of 20 points to align to the Shifts and major features of the Common

# Alignment Criterion 3

## Reading – Aligned Use of Item Types

### Metric

---

**AC Metric 3A:**

Assessments employ at least one item type that requires students to write rather than select a response (brief or extended constructed-response or performance tasks), so that the depth and complexity of the Standards can be strategically addressed.

If the above metric is met, assign the materials 2 points. If the tests employ more than one item type but do not include an item type that requires students to write a response, assign 1 point.

### Procedure for Evaluation

---

Tests that are well aligned (2 points) make use of an item type that requires writing rather than selecting a response (brief or extended constructed-response items or performance tasks). If additional item types are used, they may be selected-response in format.

Tests that are moderately aligned (1 point) do not offer constructed-response or performance tasks but make use of at least two different selected-response item types (e.g., multiple-choice, two-part evidence-based selected-response items, technology-enhanced items).

For each grade, examine the questions assembled for Non-Negotiable 2A above.

Determine which item types are being used. Note whether or not constructed-response items (either brief or extended) are included.

### Evidence

---

### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

# Alignment Criterion 3

## Reading – Aligned Use of Item Types

### Metric

---

**AC Metric 3B:**

At least 50% of the score points on each assessment are derived from items that require students to directly provide evidence from the text to support answers, i.e., the items ask students to provide details (quotations or paraphrases) from the text in support of text-based claims or inferences.

If the above metric is met, assign the materials 2 points. If 30% to 50% of score points on each assessment are derived from items that require students to directly provide textual evidence, assign 1 point.

### Procedure for Evaluation

---

Aligned tests emphasize reading and writing grounded in evidence from text, both literary and informational. Formats requiring direct use of evidence include:

- Constructed-response (CR), requiring students to use textual evidence in written responses
- Two-part evidence-based selected-response (EBSR), with one part asking for textual evidence
- Technology-enhanced (TE), requiring students to select or locate evidence within a passage
- One-part multiple-choice (MC) or TE with answer options consisting of textual details (e.g., actual quotations from the text)

For each grade, examine the questions assembled for Non-Negotiable 2A above, along with the passages on which the questions are based.

List the sequence numbers of the questions that require direct use of textual evidence.

Determine the number of score points.

Calculate percentages.

### Evidence

---

### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

# Alignment Criterion 3

## Reading – Aligned Use of Item Types

### Metric

---

#### AC Metric 3C:

Research-focused performance tasks require students to analyze, synthesize, organize, and use information from sources; such tasks comprise a significant part of the assessments—at least 8 score points per test (which is a generally accepted minimum for a reporting category).

If the above metric is met, assign the materials 2 points. If research is assessed with aligned test questions instead of performance tasks and also offers at least 8 score points, assign 1 point. Assign 0 points if fewer than 8 score points are devoted to research skills. NOTE: Many well-aligned programs place research tasks within a writing assessment. If the materials being evaluated include research tasks, and the tasks are located in the writing assessment, the 2 points should still be awarded here.

### Procedure for Evaluation

---

Aligned performance tasks are based on paired or multiple texts, and they measure one or more Standards that focus on research skills, e.g., Reading Standard 7, Reading Standard 9, Writing Standard 7.

Aligned test items are based on paired or multiple texts, and they specifically require students to analyze, synthesize, organize, and use information from sources (e.g., not merely identify a title of a likely source or a section in a table of contents).

For each grade, determine if there is sufficient coverage of research and if the questions meet this metric:

- Examine either the test blueprints or other test specifications.
- Determine whether or not there are at least 8 score points devoted to research tasks or test questions.
- Examine the questions labeled as assessing research.
- Determine whether or not the questions require analysis, synthesis, organization, and use of information rather than simple identification.

### Evidence

---

#### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

# Directions for Alignment Criterion 4

Reading – Test Blueprints and Score Reports

## Alignment Criterion 4: Test blueprints and the corresponding score reports reflect the focus of the Standards.

### 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
- 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
- Test blueprints and other explanatory material focused on test design, including sample score reports if available

Core State Standards. The more points the materials receive on the Alignment Criterion, the better the materials are aligned.

Evaluators should provide examples of evidence in support of the numerical rating for each metric, including evidence of any specific gaps in the materials.

### Rating this Criterion

Rate the assessments for each of the metrics for these Alignment Criteria as Meets (2 points), Partially Meets (1 point), or Does Not Meet (0 points). Then rate the Alignment Criteria as a group as Meets or Does Not Meet, based on the minimum number of points required. For the Alignment Criteria for reading, materials must earn at least 16 of 20 points to align to the Shifts and major features of the Common

# Alignment Criterion 4

## Reading – Test Blueprints and Score Reports

### Metric

---

**AC Metric 4A:**

Test blueprints and score reports for reading tests are based on ELA/literacy domains that are research-based and instructionally actionable (not CCSS cluster headings).

If the above metric is met, assign the materials 2 points. If a majority of the reporting categories are research-based and actionable, assign 1 point.

### Procedure for Evaluation

---

Potential reporting categories include: Reading, Writing, Reading Literature, Reading Informational Texts, Research, Vocabulary. This list is not exhaustive, and reading assessments can align to the CCSS without providing all of these categories, depending on the purpose of the test. However, such CCSS cluster headings as “Key Ideas and Details” or “Craft and Structure” are not appropriate for use as reporting categories, as they were not designed to provide research-based instructionally actionable guidance.

For each grade, examine either the test blueprints, other test specifications or sample score reports.

Determine and evaluate the names of the reporting categories and sub-categories.

### Evidence

---

### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

# Alignment Criterion 4

## Reading – Test Blueprints and Score Reports

### Metric

---

**AC Metric 4B:**

Test blueprints balance total reading word count and item counts per test form with time allotted, so that students have sufficient time and purpose to read carefully and deeply. On average, passages have 7 to 10 score points each.

If the above metric is met, assign the materials 2 points. If testing time allows for some rereading and passages have an average of 5 to 6 score points each, assign 1 point.

### Procedure for Evaluation

---

Aligned assessments do not overburden students with a large number of texts in a short time period and/or offer only a few score points for each text. Standards-based questions are designed to send students back to the text for rereading, and assessments should allow sufficient time. Also, item sets should be large and robust enough to provide an appropriate balance between the number of texts and numbers of questions, so that students are not asked to read a complex text but given only a few questions to answer.

For each grade, examine the following:

- The test form(s)
- The test blueprints
- The specifications for time allotted

Determine the ratio of passage sets to time allotted, judging if there is sufficient time for rereading the passages.

Calculate the average number of score points per passage.

### Evidence

---

#### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

# Alignment Criteria 1–4

## Reading

### Points Assigned for Alignment Criteria 1–4

---

Materials must earn at least 16 of 20 points to meet the Alignment Criteria 1 – 4 for reading. If materials earn fewer than 16 points, the Criteria have not been met. Check the final rating.

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

### Rating

---

\_\_\_\_ Total (20 points possible)

Meets

Does Not Meet

**Strengths / Weaknesses:**

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

# Directions for Alignment Criterion 5

Writing – Writing to Sources

**Alignment Criterion 5: Writing tasks reflect the writing types named in the Standards and require students to write to sources.**

## Required Materials

- The writing prompts, performance tasks, or constructed-response test questions in the test forms for each grade level or (for an item bank) a representative sample of writing prompts, performance tasks, or constructed-response test questions
- Metadata accompanying the prompts, tasks, or questions, showing the alignment of each to the CCSS
- Test blueprints and other explanatory material focused on test design, including sample score reports if available

Evaluators should provide examples of evidence in support of the numerical rating for each metric, including evidence of any specific gaps in the materials.

NOTE: Many well-aligned programs place research tasks within the reading assessment rather than the writing assessment. If the materials being evaluated include research tasks in the writing assessment, evaluate those tasks using Reading Alignment Criterion 3C.

## Rating this Criterion

Rate the assessments for each of the metrics for this Alignment Criterion as Meets (2 points), Partially Meets (1 point), or Does Not Meet (0 points). Then rate the Alignment Criterion as a whole as Meets or Does Not Meet, based on the minimum number of points required. For the Alignment Criterion for writing, materials must earn at least 3 of 4 points to align to the Shifts and major features of the Common Core State Standards. The more points the materials receive on the Alignment Criterion, the better the materials are aligned.

# Alignment Criterion 5

## Writing – Writing to Sources

### Metric

---

**AC Metric 5A:**

Taking all forms of the test together, 100% of writing tasks within a grade band approximate the balance of exposition, persuasion, and narrative required by the Standards (or blend writing types in similar proportions):

Grades 3-5:        exposition 35%  
                          opinion 30%  
                          narrative 35%

Grades 6-8:        exposition 35%  
                          argument 35%  
                          narrative 30%

High School:      exposition 40%  
                          argument 40%  
                          narrative 0-20%

If the above metric is met, assign the materials 2 points. If narrative writing is greater than the indicated percentages but less than 50% of writing tasks, assign 1 point. If narrative tasks comprise more than 50% of writing tasks, assign 0 points.

### Procedure for Evaluation

---

As students progress through the grades, an increasing focus on both argument and explanatory/informational writing is crucial for readiness.

For each grade band, examine either the writing tasks and/or constructed-response items on the Reading and Writing test form(s) or a representative sample (a minimum of 15 prompts or tasks) from the item bank for each grade band.

List the writing type for each task or item.

Calculate percentages of each of the three types of writing within each band.

### Evidence

---

### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

# Alignment Criterion 5

## Writing – Writing to Sources

### Metric

---

**AC Metric 5B:**

At least 90% of expository and argument/persuasive writing tasks require writing to sources — i.e., students confront text directly, draw on textual evidence, and support valid inferences from the text.

If the above metric is met, assign the materials 2 points. If 75% to 90% of expository and argument/persuasive prompts require writing to sources, assign 1 point.

### Procedure for Evaluation

---

For expository and argument/persuasive prompts, students should be required to read texts and draw on textual evidence to support valid claims and inferences.

For each grade band, examine the writing items from Alignment Criterion 5A above.

List any prompts or tasks that do not require writing to sources.

Calculate the percentage of expository and argument/persuasive prompts requiring writing to sources within each band.

### Evidence

---

### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

# Alignment Criterion 5

Writing – Writing to Sources

**Alignment Criterion 5: Writing tasks reflect the writing types named in the Standards and require students to write to sources.**

## Points Assigned for Alignment Criterion 5

---

Materials must earn at least 3 of 4 points to meet Alignment Criterion 5 for writing. If materials earn fewer than 3 points, the Criterion has not been met. Check the final rating.

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

## Rating

---

\_\_\_\_ Total (4 points possible)

Meets

Does Not Meet

**Strengths / Weaknesses:**

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

# Directions for Alignment Criterion 6

## Language

**Alignment Criterion 6: Test questions assessing conventions and writing strategies focus on the specifics of the Standards and reflect actual practice to the extent possible.**

### 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 questions, showing the alignment of each to the CCSS
- Test blueprints and other explanatory material focused on test design, including sample score reports if available

### Rating this Criterion

Rate the assessments for each of the metrics for this Alignment Criterion as Meets (2 points), Partially Meets (1 point), or Does Not Meet (0 points). Then rate the Alignment Criterion as a whole as Meets or Does Not Meet, based on the minimum number of points required. For the Alignment Criterion for language, materials must earn at least 4 of 6 points to align to the Shifts and major features of the Common Core State Standards. The more points the materials receive on the Alignment Criterion, the better the materials are aligned.

Evaluators should provide examples of evidence in support of the numerical rating for each metric, including evidence of any specific gaps in the materials.

# Alignment Criterion 6

## Language

### Metric

---

**AC Metric 6A:**

At least 90% of language score points are derived from questions that focus on the specifics of the language Standards for the grade, assessing common errors and skills important for readiness.

If the above metric is met, assign the materials 2 points. If at least 50% of language score points are derived from questions that assess common student errors and focus on the conventions and strategies most important for readiness, assign 1 point.

### Procedure for Evaluation

---

Questions focused on English conventions and writing strategies should represent common student errors (not artificial or unlikely mistakes).

Questions should focus on the conventions most important for college and career readiness as indicated by the Standards (see “Language Progressive Skills, by Grade” ([http://www.corestandards.org/assets/CCSSI\\_ELA%20Standards.pdf](http://www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf))).

For each grade, examine either the language questions on the test form(s) or a representative sample (at least 15 language questions per grade) in the item bank.

List the sequence numbers of the questions that do not focus on the specifics of the Standards at each grade level or do not assess common errors.

Calculate percentages.

### Evidence

---

#### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

# Alignment Criterion 6

## Language

### Metric

---

**AC Metric 6B:**

At least 60% of language score points in the test blueprints are derived from students' written responses and/or technology-enhanced items that mimic actual editing, mirroring real-world activity as closely as possible.

If the above metric is met, assign the materials 2 points. If at least 40% of language score points in the test blueprints are derived from actual writing and/or technology-enhanced items that mimic actual editing, assign 1 point.

### Procedure for Evaluation

---

Students should demonstrate language skills in the context of actual written composition, with use of conventions and writing strategies explicitly designated as part of the scoring rubric.

Alternately or in addition to actual writing, students should be asked to do editing or revision using technology-enhanced items that mimic actual editing and revision.

Using the list of items generated for Alignment Criterion 6A above, list the sequence numbers of the questions that do not mirror real-world activity.

Calculate percentages.

### Evidence

---

### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

# Alignment Criterion 6

## Language

### Metric

---

**AC Metric 6C:**

Unless only reading and writing are being assessed, language skills questions comprise a sufficient part of ELA/literacy assessments—at least 8 score points per test (which is a generally accepted minimum for a reporting category).

If the above metric is met, assign the materials 2 points. If there are at least 6 language score points, assign 1 point.

### Procedure for Evaluation

---

At each grade, each assessment should include a sufficient number of points for language skills so that language could be a reporting category. Providing a reporting category for language is desirable but is not required.

For each grade, examine either the test form(s) or the test blueprints.

Determine the number of test questions or score points devoted to language skills at each grade level.

### Evidence

---

### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

# Alignment Criterion 6

Language

**Alignment Criterion 6: Test questions assessing conventions and writing strategies focus on the specifics of the Standards and reflect actual practice to the extent possible.**

## Points Assigned for Alignment Criterion 6

---

Materials must earn at least 4 of 6 points to meet Alignment Criterion 6 for language. If materials earn fewer than 4 points, the Criterion has not been met. Check the final rating.

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

## Rating

---

\_\_\_\_ Total (6 points possible)

Meets

Does Not Meet

**Strengths / Weaknesses:**

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

# Directions for Alignment Criterion 7

## Speaking and Listening

**Alignment Criterion 7: Test questions assessing speaking and listening reflect true communication skills required for college and career readiness.**

### 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 questions, showing the alignment of each to the CCSS
- Test blueprints and other explanatory material focused on test design, including sample score reports if available

### Rating this Criterion

Rate the assessments for each of the metrics for this Alignment Criterion as Meets (2 points), Partially Meets (1 point), or Does Not Meet (0 points). Then rate the Alignment Criterion as a whole as Meets or Does Not Meet, based on the minimum number of points required. For the Alignment Criterion for speaking and listening, materials must earn at least 3 of 4 points to align to the Shifts and major features of the Common Core State Standards. The more points the materials receive on the Alignment Criterion, the better the materials are aligned.

Evaluators should provide examples of evidence in support of the numerical rating for each metric, including evidence of any specific gaps in the materials.

# Alignment Criterion 7

## Speaking and Listening

### Metric

---

**AC Metric 7A:**

When speaking is being assessed, at least 75% of the test questions require active speaking tasks rather than selected-response or technology-enhanced items about speaking practices.

If the above metric is met, assign the materials 2 points. If at least 50% of the test questions require active speaking tasks, assign 1 point.

### Procedure for Evaluation

---

Questions assessing speaking focus on students' ability to engage effectively in a range of conversations and collaborations. Students should be asked to express and support ideas clearly and effectively, probing ideas under discussion by building on others' ideas.

For each grade, examine either the speaking questions on the test form(s) or a representative sample (at least 15 speaking questions per grade) in the item bank.

List the sequence numbers of the questions that do not focus on the skills required for readiness and require active speaking.

Calculate percentages.

### Evidence

---

### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

# Alignment Criterion 7

## Speaking and Listening

### Metric

---

**AC Metric 7B:**

When listening is being assessed, at least 75% of the test questions require active listening rather than selected-response or technology-enhanced items about listening practices.

If the above metric is met, assign the materials 2 points. If at least 50% of the test questions require active listening skills, assign 1 point.

### Procedure for Evaluation

---

Students should be asked to express and support ideas clearly and effectively, probing ideas under discussion by building on others' ideas.

Students should also be asked to demonstrate such skills as taking notes on main ideas and asking relevant questions.

For each grade, examine either the listening questions on the test form(s) or a representative sample (at least 15 listening questions per grade) in the item bank.

List the sequence numbers of the questions that do not focus on the skills required for readiness and require active listening.

Calculate percentages.

### Evidence

---

### Rating

---

- Meets (2)
- Partially Meets (1)
- Does Not Meet (0)

# Alignment Criterion 7

Speaking and Listening

**Alignment Criterion 7: Test questions assessing speaking and listening reflect true communication skills required for college and career readiness.**

## Points Assigned for Alignment Criterion 7

---

If the assessments include both speaking and listening, materials must earn at least 3 of 4 points to meet the Alignment Criterion for speaking and listening. If materials earn fewer than 3 points, the criterion has not been met.

If the assessments include either speaking or listening, materials must earn at least 1 point to meet the Alignment Criterion. If materials do not receive at least 1 point, the Criterion has not been met. Check the final rating.

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

## Rating

---

\_\_\_\_ Total (4 or 1 points possible)

Meets

Does Not Meet

**Strengths / Weaknesses:**

**Move to the Evaluation Summary on the following page to record the final Meets or Does Not Meet rating.**

# AET Evaluation Summary 1 of 2

ELA/Literacy, Grades 3–12

Assessment Evaluation Tool (AET)

ELA/Literacy, Grades 3 -12

Title of Assessment: \_\_\_\_\_

Name of Evaluator(s): \_\_\_\_\_

Publisher: \_\_\_\_\_

Date of Evaluation: \_\_\_\_\_

Signature of Each Evaluator(s): \_\_\_\_\_

## Non-Negotiable Alignment Criteria

Each Non-Negotiable must be met in order for the Non-Negotiable Alignment Criteria to be met overall. The Non-Negotiables apply to Reading assessments. If Reading is not intended to be part of the assessment, indicate N/A.

### Non-Negotiable 1: Complexity of Texts

Meets  N/A  
 Does Not Meet

### Non-Negotiable 2: Text-Dependent and Standards-Based Questions

Meets  N/A  
 Does Not Meet

### Non-Negotiable Overall:

Meets  
 Does Not Meet

## Alignment Criteria

Each Alignment Criterion relevant to the assessments evaluated must be met with a sufficient number of points in order for the Alignment Criteria to be labeled as “Meets” overall. If a particular domain is not intended to be part of the assessment, indicate N/A for that criterion. The more points the materials receive on the relevant alignment criteria, the better they are aligned.

### Alignment Criteria 1-4: Reading

Points: \_\_\_\_ of 20 possible.

(Materials must receive at least 16 of 20 points to align.)

Meets  N/A  
 Does Not Meet

### Alignment Criterion 5: Writing

Points: \_\_\_\_ of 6 possible.

(Materials must receive at least 3 of 4 points to align.)

Meets  N/A  
 Does Not Meet

### Alignment Criterion 6: Language

Points: \_\_\_\_ of 6 possible.

(Materials must receive at least 4 of 6 points to align.)

Meets  N/A  
 Does Not Meet

### Alignment Criterion 7: Speaking and Listening

Points: \_\_\_\_ of 4 possible.

(Materials that include both speaking and listening must receive at least 3 of 4 points to align; materials that assess either speaking or listening must receive at least 1 point.)

Meets  N/A  
 Does Not Meet

### Alignment Criteria Overall:

Meets  
 Does Not Meet

# AET Evaluation Summary 2 of 2

ELA/Literacy, Grades 3–12

Assessment Evaluation Tool (AET)

ELA/Literacy, Grades 3 -12

Title of Assessment: \_\_\_\_\_

Name of Evaluator(s): \_\_\_\_\_

Publisher: \_\_\_\_\_

Date of Evaluation: \_\_\_\_\_

Signature of Each Evaluator(s): \_\_\_\_\_

## Summary

---

If the materials meet both Non-Negotiables and relevant Alignment Criteria, they are aligned to the Shifts and major features of the CCSS.

**Do the materials meet both Non-Negotiables and relevant Alignment Criteria?**

Yes

No

**What are the specific areas of strength and weakness based on this review?**

Publishers or others modifying or developing assessments can use this information to make improvements and/or to remedy gaps in the alignment of assessment materials.

# Indicators of Quality

Once an evaluation for alignment to the Shifts and major features of the CCSS has been conducted using Sections 1 – 3, it is important to evaluate for overall quality and best practices. A starting list of Indicators of Quality is suggested below, including critical considerations such as accessibility for all students. States, districts and others evaluating assessment options are encouraged to add to this list to ensure materials respect curricular choices and reflect local contexts. These indicators are designed to apply to assessment programs; and similar indicators are reproduced in the Quality Criteria Checklists, which are used to evaluate individual passages and test questions.

## Indicators of Quality for Assessment Programs

---

## Evidence

---

1. Assessments must provide accessibility to all students, including English learners and students with disabilities: The assessments should be 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. Allowable accommodations and modifications that maintain the constructs being assessed should be offered where appropriate.
2. Assessments should indicate progress toward college and career readiness: Scores and performance levels on assessments should be 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.
3. Assessments must be valid for required and intended purposes. As appropriate, assessments produce data, including student achievement data and student growth data, that can be used to validly inform individual student gains and performance and other purposes such as school effectiveness and improvement.
4. Assessments must be reliable: Assessments minimize error that may distort interpretations of results, describe the precision of the assessments at the cut scores, and are generalizable for the intended purposes.
5. Assessments should be 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), and score scales used facilitate accurate and meaningful inferences about test performance.

# Indicators of Quality

## Indicators of Quality for Texts Used in Reading Assessments

---

## Evidence

---

1. Excerpts should convey a sense of completeness: When texts are excerpts from a larger work, they should begin and end logically and maintain the intent of the original. Edits for length should be made at the beginning or end of the piece, rather than in patchwork fashion.
2. Introductory material should include only the most necessary information. When the texts are presented with introductory material, the introduction should avoid summarizing or explaining the meaning of the text or giving students answers to questions.
3. Illustrations should add value. When texts include visual elements, the elements should be related to the central ideas of the text and provide important additional information.
4. Texts should fall within an acceptable range of word count. All texts should fall within an acceptable range for word count at the indicated grade level.
5. Paired or multiple texts should have a clear and meaningful relationship with each other. When texts are paired, the potential points of comparison should be significant (not superficial), such as theme, amount and quality of evidence, differences in emphasis, distinguishable structures, changes to derivative text.
6. For tasks that simulate research, one text should serve as an “anchor” text. When research tasks are presented, the first text in the set should provide foundational knowledge and lead naturally to additional reading and exploration.

# Indicators of Quality

## Indicators of Quality for Test Questions

---

## Evidence

---

1. The language used in Reading items and Writing prompts should be clear and concise. The language in the items should reflect vocabulary and sentence structures appropriate to the grade level.
2. Selected-response items should be presented for review with rationales for all answer choices. The metadata for the selected-response items (multiple-choice or technology-enhanced) should provide a rationale for every answer option.
3. Selected-response items should exemplify high standards of technical quality. If items use a selected-response format, they should be free from internal clueing (e.g., the options should not repeat words in the stem; the grammatical relationship between stem and options should be correct for all options, the correct response should not be more specific than the options, the correct answer should not simply paraphrase words in the text). Also, the distractors should be plausible but incorrect (not unintended or arguable correct answers); general statements (e.g., central idea, theme, structure) should be precise and accurate; and inferences should be provable with specific textual evidence.
4. Constructed-response items should be presented for review with sample responses. The metadata for constructed-response items (brief and/or extended response) should provide a top score response or a sample response for every score point.
5. Constructed-response items should exemplify high standards of technical quality. If the items ask students to generate a written response, the description of the task should be clear enough that students know the characteristics of a successful response. Also, items that ask for a written response should be accompanied by information for students about the criteria for scoring.
6. Two-part items should exemplify high standards of technical quality. If items have two parts, the relationship between the two parts should be clear and logical, and there should be a plausible link between the options in the two parts.
7. Technology-enhanced items should exemplify high standards of technical quality. If items use computer delivery, they should use technology to approach the text in ways other item types cannot, providing value beyond that of non-technology-enhanced items. Also, the directions for use of technology should be clear and easy to follow.

# Indicators of Quality

## Indicators of Quality for Test Questions

---

8. Items that call for comparison or synthesis should focus on meaningful aspects of the texts. Questions that ask for comparison or synthesis should be related to central (rather than trivial) aspects of the text (e.g., amount and quality of evidence, differences in emphasis, distinguishable structures, changes to derivative text).
9. Graphic organizers used in items should be text-specific and add value. When items have graphic organizers or similar formats, the organizer should arise from characteristics of the text, i.e., it should not be a generic format that could apply to any text. The organizer or format should add value to the item by allowing students to demonstrate understanding of the text in a way that a traditional item would not.

## Evidence

---

## Indicators of Quality for Sets of Test Questions

---

1. As a whole, a set of items should allow students to demonstrate deep understanding of the text. Sets of items should require students to read the full text carefully and show their understanding of the central ideas, allowing and requiring students to provide deep insights rather than skim the surface.
2. As a whole, a set of items should cover the Standards that arise naturally from the unique aspects of the text. Sets of items should address as many different Standards as appropriate, with items based on the individual characteristics of the texts rather than on a forced standard coverage design.
3. As a whole, a set of items should be ordered in a logical and helpful manner (unless item order cannot be fixed, i.e., the items are delivered in an adaptive system or are collected in an item bank). Sets of items should begin and/or end with general questions about the text; questions about particulars of the text should be presented in the order the particulars appear in the text.

## Evidence

---