Quality Criteria for Assessment Items: Questions Worth Answering

The following criteria are designed to help item reviewers determine if an item or set of items align(s) to the Common Core State Standards (CCSS). The criteria are set up in a gated manner so that reviewers can quickly and systematically determine where the item or set of items strays from the expectations of the CCSS.

Review the item or set of items against the criteria in order, and place a "Y" or an "N" in the middle column, labeled "Y/N". Please use the "Explanation" column to clarify the response or recommend a revision when an item or an item set receives an "N." If a criterion does not apply to a particular item, leave the column blank.

There are four gates in the checklist. The first two gates pertain to all items individually. The third and fourth gates apply to sets of items. If an item does not meet the criteria in the first gate, it should be removed from consideration. If it does meet the criteria in the first gate, review it according to the criteria in the second gate. The item must meet or must be able to be revised to meet the criteria in the second gate. At the end of the second gate, recommend whether to accept, accept conditionally, or reject the item.

Sets of items must meet the criteria in the third gate, and they should be revised to meet relevant criteria in the fourth gate. Again, reviewers will be asked to make a recommendation whether to accept, accept conditionally, or reject the item sets.

Criteria for Evaluating Items for CCSS-Aligned Assessments		
1. FIRST GATE: SINGLE ITEMS – The item must meet all of the criteria in the first gate to be considered further.	**Y/N**	**Explanation**
1.A Value: Is the item worthy of student attention, and does it allow students to deliver insights about the text?		
1.B Text Dependency: 1. Does the item require close analytic reading of the text (either close reading of part of a text or the entire text)? Providing the correct answer should not require prior knowledge, nor should it be possible for students to answer the question without reading the text. 2. Does the item require students to use evidence from the text either by directly asking students cite evidence or by requiring students to use evidence to provide the answer? An item should require students to follow the details of what is explicitly stated and/or make valid inferences.		
1.C Alignment: Does the item clearly align with the intent and language of one or more Common Core State Standard(s) or evidence statement(s)/target(s), including Reading standard 1? • If the item has a different alignment from the one(s) indicated, write a "Y" in the middle column and give details about a proposed change in alignment in the Explanation column.		
1.D Rationales and/or Top-Score Responses: For an SR item, are effective rationales, which describe the answer choices rather than predict student behavior, provided for the correct answer and each distractor? For a CR item, are sample responses provided for each score point?		

If the item does not have a "Y" in all of the criteria above, remove the item from consideration. If the item does meet the criteria in the first gate, proceed to the second gate for single items.			
2.	**SECOND GATE: SINGLE ITEMS -** Items that pass the first gate must meet or be revised to meet the following, as applicable:.	**Y/N**	**Explanation**
2.A	**Text Specificity:** Is the item not only text-dependent but also text-specific—not a generic question, but one that arises organically from the text and applies the language of the standards as appropriate to the text? • If yes, place a "Y" in the column on the right. • If the item is not text-specific, place an "N" in the column on the right and either suggest a revision or give reasons for keeping the item as is.		
2.B	**Clarity of Language:** Is the language used in the item clear and concise, and does it avoid negative phrasings and complex sentence structures (unless such structures are being tested)? • If yes, place a "Y" in the column to the right. • If the item should be revised for clarity of language, place an "N" in the column at the right and specify problems and/or suggest revisions in the Explanation column.		
2.C	**Technical Quality:** Does the item exemplify high standards of technical quality? For an SR item, for example, the question precludes guessing (plausible distractors or gridded response), the correct response is defensible based on textual evidence, no option is conspicuous and therefore possibly inviting, etc. For a CR item, for example, there is a clear description of the task and the criteria for scoring. *The above descriptions of technical quality are not exhaustive; reviewers should call on their knowledge of all best practices to evaluate technical quality.* • If there are no concerns about technical quality, place a "Y" in the column to the right. • If there are concerns, place an "N" in the column at the right and specify problems and/or suggest revisions in the Explanation column.		
2.D	**Technology:** If technology is used: Does it provide value beyond that of a non-technology-enhanced item (i.e., no use of technology for technology's sake, no confusing instructions or complicated actions)? And does the technology avoid introducing a new construct other than close reading and use of evidence—a construct that is not required by the CCSS? • If yes, place a "Y" in the column to the right. • If the use of technology should be improved or eliminated, place an "N" in the column at the right and detail the concerns in the Explanation column.		

2.E	**Hand Scoring:** If the item is to be hand-scored, does it provide information beyond what would be gained from a selected-response or machine-scored item? • If yes, place a "Y" in the column to the right. • If no, place an "N" in the column at the right and specify problems and/or suggest revisions in the Explanation column.		
2.F	**Comparison Items:** If the item calls for comparison or synthesis of ideas, is the comparison or synthesis meaningful and related to central ideas in the text? • If yes, place a "Y" in the column to the right. • If no, place an "N" in the column at the right and specify problems and/or suggest revisions in the Explanation column.		
2.G	**Graphic Organizers:** If the item contains a graphic organizer or similar format, does the organizer or format add significant value to the item by allowing students to demonstrate knowledge in a way that a traditional selected-response item would not? (Use of graphic organizers or other narrow formats in test items may tend to change the construct being tested or to privilege these devices over others and thus influence teachers to include them in instruction.) • If the organizer or format adds value, assesses the construct of reading, and is text-specific so that it is not likely to solidify in instruction, place a "Y" in the column to the right. • If the organizer does not add value or risks changing the construct or solidifying in instruction, place an "N" in the column at the right and specify problems and/or suggest revisions in the Explanation column.		
2.H	**Vocabulary Items:** If the item assesses vocabulary, does it focus on crucial academic (tier 2) vocabulary in context and do the distractors reflect the same part of speech as the word being tested? • If yes, place a "Y" in the column to the right. • If the item tests a non-tier 2 word or tests other vocabulary skills besides use of context, place an "N" in the column at the right. Specify problems, suggest revisions, or give reasons that justify retaining the item.		
Accepted (all "Y's") Accepted conditionally, with comments to be addressed Rejected		☐ ☐ ☐	

3.	**THIRD GATE: ITEM SETS (ITEMS ASSOCIATED WITH A TEXT OR TEXTS) –** Items in a set must pass the first two gates individually. Item sets must then meet the criterion in this gate to be considered further.	**Y/N**	**Comments**

3.A **Comprehensiveness:** Does the set require students to read the full text carefully and show their understanding of the central ideas in the text (the set allows and requires students to provide read for deep insights rather than skim the surface)?		
If the item set does not have a "Y" for the criterion above, remove the set from consideration. If the set does meet the above criterion, proceed to the fourth gate below.		
4. **FOURTH GATE: ITEM SETS (ITEMS ASSOCIATED WITH A TEXT OR TEXTS) –** Item sets that pass the third gate must meet or be revised to meet the following criteria in this gate, as applicable.	**Y/N**	**Comments**
4.A **Standard Coverage:** Does the set address as many different Standards (and evidence statements/targets) as possible, with items based on the individual characteristics of the text and focused on key aspects of the text? The set of items should be extensive and robust enough that a good selection of items will remain after field testing. • If yes, place a "Y" in the column to the right. • If no, place an "N" in the column at the right and give reasons in the Explanation column.		
4.B **Item Cluing:** Do the items avoid cluing the answer to other items in the set? • If yes, place a "Y" in the column to the right. • If no, place an "N" in the column at the right and indicate in the Explanation column which items clue each other so that they can be marked in the bank appropriately (not to be used on the same form).		
Accepted (all "Y's") Accepted conditionally, with comments to be addressed Rejected	☐ ☐ ☐	