CSAI Report

Mathematics Item Writing Guide and Item Exemplars

June 2016

The work reported herein was supported by grant number #S283B050022A between the U.S. Department of Education and WestEd with a subcontract to the National Center for Research on Evaluation, Standards, and Student Testing (CRESST). The findings and opinions expressed in this publication are those of the authors and do not necessarily reflect the positions or policies of CRESST, WestEd, or the U.S. Department of Education.

WestEd is a nonpartisan, nonprofit research, development, and service agency that works with education and other communities throughout the United States and abroad to promote excellence, achieve equity, and improve learning for children, youth, and adults. WestEd has more than a dozen offices nationwide, from Massachusetts, Vermont and Georgia, to Illinois, Arizona and California, with headquarters in San Francisco. For more information about WestEd, visit WestEd.org; call 415.565.3000 or, toll-free, (877) 4-WestEd; or write: WestEd / 730 Harrison Street / San Francisco, CA 94107-1242.

Table of Contents

Introduction

Purpose of This Guide

Although instructional staff are already experienced in the creation and use of classroom assessments, there is a demand for in-depth guidance and professional development related to assessment literacy. This guide is intended to support professional learning related to assessment design and assessment literacy, with a particular focus on formative assessment. The guide includes detailed explanations of assessment-related concepts that are relevant to the needs of instructional staff. Item exemplars from across the K–HS grades illustrate how these concepts are applied, and commentary on each item provides further insight into the development of high-quality mathematics assessment items.

The ultimate goal of this guide is to help instructional staff refine their senses of what constitutes a high-quality mathematics assessment item, and to understand, in some detail, reasons that a particular item is or is not successful. Staff can apply this knowledge to create high-quality formative assessments for classroom use, and/or to develop local assessments for use in meeting accountability and evaluation requirements. Enhanced assessment literacy can also inform the evaluation of commercial assessments that are being considered for use at the local and/or state levels.

How to Use This Guide

This guide should be read from beginning to end, since the sections build on one another to demonstrate the craft and principles of item development.

The first part of this guide contains three sections that, together, provide an overview of the principles of sound item development. The first section discusses common item formats and their component parts in detail, with a particular emphasis on the mechanics of constructing high-quality multiple-choice (also referred to as selected-response) and constructed-response items. Readers will gain familiarity with the building blocks with which item writers work—and with how these parts are assembled into a successful whole.

These foundational principles of item writing provide the basis for the second section, which focuses on developing well-aligned, rigorous, and grade-appropriate mathematics content and language. In addition to alignment, this section considers rigor and depth of knowledge, and distinguishes between rigor and complication; it also pays special attention to accessibility and the needs of English language learners. Readers will learn strategies for increasing the rigor of items while reducing irrelevant sources of challenge that interfere with a student's ability to demonstrate skills or knowledge.

The third section includes two useful tools: a general checklist for item development and a set of guiding questions to inform high-quality item writing.

The second part of the guide contains eleven annotated item exemplars that demonstrate the principles set forth in the guide's first part. Cross-references are provided throughout the text of the first part of the guide, so that readers can locate relevant examples among these item exemplars. One item is included from each of the grades K–8 and two items are included from the high school (HS) grade band. The item exemplars represent a variety of formats and approaches and align to a range of standards.

Each item exemplar is designed to illustrate common issues that arise during item development. In some cases, two versions of the item are included: one version at a specified depth of knowledge (DOK) level, and one version that illustrates how the item might be modified to change the DOK level. Commentary about each item provides further insight into the process of item development.

Item Development Basics: Format and Structure

Item Formats

This guide focuses on two basic item formats that are the backbone of formative assessments: multiple-choice (MC) items and constructed-response (CR) items. Innovative item formats such as technology-enhanced items (TEIs) are discussed briefly in this section.

Several different types of MC and CR items are appropriate for assessments. MC items may be single select or multiple select. CR items may require a short response of several sentences and/or a demonstration of a mathematical procedure, or may require a longer response in the form of an essay explaining mathematical reasoning, sometimes including the construction of figures or graphs. TEIs may involve tasks such as matching, sorting, dragging and dropping answers into a table or chart, or using drop-down menus to complete sentences or equations.

Table 1 describes the characteristics of each item format and provides rationales for using each format.

Table 1—Item Format Overview

Item Format	**Characteristics**	**Rationale for Use**
Single-select multiple choice (SSMC)	One response is selected from among several answer choices The industry-standard number of choices is four	Versatile, user-friendly format that can measure a range of constructs, including higher-order thinking skills Use when there is one and only one correct answer, or when one answer is clearly the strongest possibility

Multiple-select multiple choice (MSMC)	Two or more responses are selected from among several answer choices Desired number of responses is sometimes indicated in the stem Advisable to have more distractors (incorrect answers) than correct answers (e.g., four distractors and two correct answers, for a total of six answer choices) Partial credit may be awarded for selecting some, but not all, correct answers	The construct must justify the use of a more complex item format Use when there are two or more distinct and equally strong correct answers
Constructed response (CR)	Student provides a written response to a prompt Answer may be short (e.g., one or two sentences) or long (e.g., a multi-paragraph response explaining mathematical reasoning) Scoring must be based on an item-appropriate rubric	Use for multi-step tasks that cannot be accommodated in other item formats Use to elicit a range of student responses Use to have students explain their thinking
Technology-enhanced item (TEI)	Student uses an interactive interface to perform a task or provide a response Types of TEIs include sorting, sequencing, completing tables or charts, completing sentences, and graphing Scoring is based on the number of correct interactions Partial credit should be available	Use for tasks that cannot be accomplished via a multiple-choice format Avoid TEIs that could be presented as multiple-choice items

In addition, some item formats are generally avoided on high-quality assessments because of their inherent limitations in providing information about what a student knows or can do.

Table 2 describes item formats that should be avoided, and provides explanations of their limitations.

Table 2—Item Formats to Avoid

Item Format	Limitations
True/False	Often measures lower-order thinking skills (memorization; recall) Student has a 50 percent chance of guessing the correct answer Provides no information about misconceptions that led to an incorrect response
MC with "all of the above" or "none of the above" as an answer choice	This answer choice may be perceived as a throwaway answer choice Item may provide little information about misconceptions that led to an incorrect response Use of "all of the above" may reward students for partial knowledge (if student realizes that two answer choices are correct, student may realize that "all of the above" must be correct) Cannot use "all of the above" in items that ask for the best answer Use of "none of the above" may reduce an item's ability to indicate differences between students with full knowledge and those with misinformation

Multiple-Choice Items

Single-select multiple-choice (SSMC) items:

The SSMC item is perhaps the most widely used item format. Variations are possible, but it is most common for a student to be asked to select a single correct response from among four possible responses (i.e., the answer choices consist of one correct response and three incorrect responses).

SSMC items must have one and only one correct answer. If the item has two or more potentially correct answers, the item stem and/or answer choices must be revised to ensure that there is only one defensible correct answer. If both answers are equally strong, the writer may also consider using the multiple-select multiple-choice format.

Multiple-select multiple-choice (MSMC) items: The MSMC format should be used if an item has two or more correct answers. Mathematics MSMC items most commonly have five to seven answer choices and two to five correct answers. An MSMC item with two correct answers might have three or (ideally) four distractors (incorrect answers), while an MSMC item with three correct answers might have four or five distractors. Exceeding five distractors is likely to result in an unmanageably long item, and in any case, writing more than five strong distractors is often difficult.

The MSMC format should be reserved for situations in which it is truly warranted. All correct answers should be meaningfully distinct (e.g., 0.5 and $\frac{1}{2}$ should not be separate correct answers unless fractional and decimal equivalents are the basis of the standard being assessed). (See the grade 2 item exemplar on page 26.) Avoid situations in which one correct answer is noticeably stronger or weaker than the other(s), and avoid situations in which the correct answers convey similar information or ideas. In addition, the material must support the chosen number of distractors. All of the distractors must be based on plausible student misconceptions. (See the Distractors section on page 15.)

If the material does not yield enough satisfactory correct answers and distractors to fulfill the requirements of an MSMC item, it is best for the item to use the SSMC format and to have one strong correct answer and three strong, plausible distractors.

Structure and wording of item stems: As a best practice, multiple-choice items should have closed stems. Most multiple-choice item stems should be phrased as questions, although it can sometimes be appropriate to phrase the stem as a command.

Because they are potentially challenging for English language learners, open stems should be used only in the rare instances when a closed stem would be confusing or too wordy. (See the Accessible Language section on page 16.)

Above all, the item stem should clearly convey the task that the student is required to perform. To demonstrate strong alignment, the item stem may use language from the standard. However, the item writer's primary consideration should be whether the task is presented clearly and in a grade-appropriate manner. (See the Alignment section on page 12.)

While it is sometimes desirable to phrase an item stem as a single sentence or question, most item stems can be broken into multiple sentences. This is particularly helpful for English language learners.

Use of "best," "most," and other qualifiers: When other accurate responses may be possible, it is common practice to include qualifying language in the stem so that the student is instructed to

select the best possible answer. Common qualifiers include, but are not limited to, *best, main, mainly, most,* and *most likely*.

This type of wording helps to ensure that strong distractors are not keyable. In the absence of language instructing the student to choose the best answer, a distractor that is partially correct might seem like a reasonable correct answer.

Wording of answer choices: Like item stems, answer choices should be clear and concise. The correct answer(s) should be the only keyable option(s).

It is also important for answer choices to be grammatically and logically compatible with the stem. For example, in an item that asks the student to identify a value and the reasoning for why the value is correct, the answer choices should be complete statements rather than grammatically incomplete phrases.

Answer-choice balance: It is important to ensure that answer choices are visually and structurally parallel and that no one answer choice stands out from the others. (See the Cueing and Clanging section on page 17.)

Common causes of outliers include negative constructions, inconsistent grammar, and unusual length. Answer choices that are strictly numerical or symbolic can also be outliers if they do not fit the pattern of the other choices (e.g., one negative number vs. three positive numbers, one fraction vs. three whole numbers, one three-digit value vs. three one-digit values). An answer choice that is a mathematical expression should not be the only choice to contain (or to not contain) a given variable or other type of mathematical symbol.

If it is not possible to make all answer choices grammatically, numerically, symbolically, visually, or structurally parallel, try to ensure that pairs or other groupings of answer choices are parallel (e.g., two answer choices in a single-select MC item are fractions and the other two are whole numbers).

Answer-choice order: Answer choices that are strictly numerical are usually ordered from least value to greatest value. Answer choices that are single word choices are usually ordered alphabetically or by length. For other types of answer choices, it is considered a best practice to order answer choices by length (either short-to-long or long-to-short), even when the answer choices are simply mathematical expressions or equations.

Answer choices that reflect options in a table or graph should be arranged by order of appearance in the table or graph.

Rationales: For mathematics MC items that will be reviewed by others (e.g., review committee) prior to administration, it is common practice include short rationales explaining the reasoning error or computational mistake that led to each distractor. Each distractor should reflect a common error that a student might make in answering a question, such as using the wrong order of operations, failing to completely factor a number, mistaking an interim step in a solution for the final step, etc.

One challenge in constructing the answer choices for MC items is that a common mathematical error often leads to a value that is significantly different in magnitude from the correct response. Ideally, in such cases, at least one other common error that also leads to a significantly different value should be used, in order to avoid creating an outlier.

Constructed-Response Items

Possible formats: The category of constructed-response (CR) items covers both short and extended tasks. Short-answer items may require a response of a few words, a couple of sentences, or a numerical answer with a supporting mathematical procedure, depending on the content that is addressed and the specifications of the assessment.

Because short-answer items are relatively simple to develop and score, this section focuses on longer tasks, variously known as extended-response or open-response items. (See the grade 4 item exemplar on page 31 for a discussion of a short-answer item.)

Constructed-response prompts: As with an MC stem, a successful CR prompt clearly and concisely sets up the task and conveys what is expected of the student.

Depending on the complexity of the mathematical task, some scaffolding may be incorporated into the prompt, to guide the student's thinking. However, the scaffolding should not place unnecessary restrictions on possible responses, nor should it be so extensive as to provide the student with a ready-made outline.

Constructed-response prompts for mathematics often ask the student to explain reasoning or justify a response. In some cases, these goals are most easily satisfied by prompting the student to show the mathematical procedure carried out in reaching a solution. If numerical computations are to be accepted for a constructed-response task, it is often clearest to directly indicate this expectation in the wording of the prompt (e.g., by using "Show your work" rather than "Explain your reasoning").

Scope of task: A good CR prompt should present a task that is substantial enough to merit a response of two or more sentences. The task's scope can be tested by writing an exemplary response, or by compiling a list of information that an exemplary response should include. If an item

writer finds it challenging to craft a robust response to a prompt, students will likely encounter the same difficulty, and the task likely needs to be revised to make it sufficiently complex.

At the same time, the task for a constructed-response item should be feasible for an average student. Consider the amount of time that students will have to respond to the prompt, including time for planning. If it seems unlikely that an average student will fully complete the task in the allotted time, the task should be streamlined.

A successful CR prompt will allow both higher- and lower-performing students to demonstrate what they know and can do. Higher-performing students should be able to provide substantial, clear and complete responses, while lower-performing students should at least be able to access the prompt and provide a response that addresses some elements of the mathematical reasoning required to answer the prompt. A CR prompt that is too simple or too challenging will not elicit a meaningful range of student responses.

Rubrics: All CR items must be scored using a rubric that describes the characteristics of responses at each score point. Though it is likely not a student-facing component, the rubric should be considered an essential part of a CR item.

The use of rubrics helps to document what the task requires, ensures consistency in scoring, and delineates the distinctions among score points. Though generic rubrics are sometimes used, it is preferable to create a customized rubric for each CR item. Generic rubrics may include material that is not relevant to a particular CR item, or may neglect to address essential features.

With that said, rubrics need not be overly complicated, and need not exhaustively document the features of every conceivable student response. Instead, the rubric should provide a stable framework for evaluating student responses.

Shorter CR items can often be scored using a simple two-point rubric. Full credit is awarded for responses that are correct and complete; partial credit is awarded for responses that are partially correct or partially complete; no credit is awarded for responses that are wholly incorrect, irrelevant, blank, or too brief to evaluate.

A longer, more complex CR item requires a more complex rubric. Rubrics for items with higher point values should make finer distinctions among score points. For example, a four-point rubric may award full credit for a response that is not only accurate and complete, but is also strategically organized and sophisticated in reasoning. Three points may be awarded for a response that is considered adequate—generally correct and complete, but lacking in the thoughtful explanation or justification of a top-score response. Two points may be awarded for a response that is not wholly adequate to the task but conveys some accurate mathematical reasoning, whereas a response that demonstrates minimal understanding or engagement would receive one point. Again, a zero-point

response would be a response that is wholly incorrect, irrelevant, blank, or too brief to evaluate. (See the grade HS—Statistics item exemplar on page 46.)

Technology-Enhanced Items

Technology-enhanced items (TEIs) are designed for online administration and are intended to allow students to respond in ways that are not possible via traditional paper-and-pencil administration. If a TEI can be written as an MC item, the use of the TEI format may not be justified. For example, a TEI in which the student drags and drops a single correct answer into a box would be better presented as a standard MC item.

Examples of tasks that are well suited for TEI functionalities include selecting numbers or expressions that apply to a task (e.g., selecting all prime factors for a number); using drop-down menus to select answers for a multi-task item; sequencing and ordering numbers; and categorizing and sorting mathematical expressions or numbers.

A tremendous amount of variety exists within the category of TEIs, and the possibilities of this item format largely depend on the authoring platform that is used. However, the following best practices are relevant to the creation of nearly all TEIs:

Be aware of the scope of the task.

Consider the number of interactions (actions that students are required to perform on the item) in each TEI, and weigh the number of interactions against the item's point value and the expected time that the task will take. If the point value of a TEI is the same as that of an MC item, three to six interactions may be sufficient, depending on the complexity of each interaction.

Ensure that partial credit is available.

One advantage of TEIs is that partial credit can be available. In order for this to be the case, TEIs need to have at least two interactions. A TEI with only one interaction results in "all or nothing" scoring and can likely work as an SSMC item instead; it also does not fully take advantage of this item type's available functionality.

Ensure that items can be scored.

TEIs are almost always scored automatically. Therefore, asking students to carry out a task in which there is a wide range of possible correct answers would usually not be viable for a TEI. In some cases, this problem can be addressed by using a different TEI format that allows for a similar task to be completed but allows fewer ways in which the item can be answered correctly.

Consider using graphics, if appropriate. Depending on the authoring system, it may be possible for TEIs to incorporate simple graphics. This capability can be especially useful at lower grades and with items that require static graphics, such as tables, charts, or graphs. Some TEIs might even ask students to populate a graphic organizer with preset numbers or mathematical expressions.

Be aware of how items appear to students. A TEI that is difficult to navigate or view will likely introduce unnecessary demands on students and interfere with their abilities to demonstrate the knowledge, skills, and abilities that the item is attempting to assess. TEIs should measure the student's knowledge of the tested material, not the student's ability to interact with technology. (See the grade 3 item exemplar on page 28.)

Item Development Basics: Content and Language

Alignment

To the extent possible in a particular item format, items must robustly address the standard(s) to which they are aligned.

Because some standards are more complex than others, it is difficult to generalize about how much of the material within a standard must be addressed in order to achieve a strong item alignment. Writers must often use their professional judgment to discern the essential skill or knowledge that a standard is targeting, and must then evaluate whether an item allows for an adequate demonstration of that skill or knowledge.

It is usually simple to align an item to a standard that deals with one straightforward skill (e.g., factoring an equation). An item aligns to this standard if the student is required to factor an equation; if the item does not meet this requirement, it does not align. If a standard is this straightforward, it must be addressed in its entirety in order to achieve a strong item alignment.

On the other hand, a standard may deal with a multilayered set of related skills (e.g., factoring an equation and determining the number of distinct values that are solutions of the equation). While it may be possible to craft a multiple-choice item that encompasses both components of the standard, the standard may be broken down: an item may be considered to align if it requires the student to factor an equation or if it requires the student to determine the number of distinct values that are solutions of an equation. (Note that, as this example shows, it may be necessary to carry out the first component of the standard in order to answer an item that focuses on the second component, though completion of the first component need not be demonstrated.)

Particularly within the Common Core State Standards (CCSS), a standard may include multiple components that simply cannot be assessed together in a cohesive and well-designed item. In this case, an item must focus on a portion of the standard. This approach does not compromise the strength of the item's alignment to the standard.

For example, geometry items at grade 6 may address the conceptual correlation of the volume formula for right rectangular prisms that have fractional edge lengths with the concept of packing

the prism with unit cubes that have unit fractional edge lengths, **or** may address application of the volume formula to solve real-world and mathematical problems:

Find the volume of a right rectangular prism with fractional edge lengths by packing it with unit cubes of the appropriate unit fraction edge lengths, and show that the volume is the same as would be found by multiplying the edge lengths of the prism; apply the formulas $V = lwh$ *and* $V = bh$ *to find volumes of right rectangular prisms with fractional edge lengths in the context of solving real-world and mathematical problems.*

To achieve strong alignment, the language of the standard can be incorporated into the wording of the item. For example, the stem of an item aligned to this standard might read:

What is the volume, in cubic meters, of a right rectangular prism with edge lengths of $5\frac{3}{4}$ *m,* $2\frac{1}{2}$ *m, and* $3\frac{1}{8}$ *m?*

However, the language of the standard should not be used verbatim if doing so results in an item that is unclear or inaccessible for the targeted grade level. In such cases, items should be worded in a way that captures the essence of the standard, even if the exact language differs. While using the language of the standard can support strong alignment, it is more important to consider the essential skills and knowledge that the standard is targeting, and to craft the item in a way that presents the task clearly and at a grade-appropriate level.

Furthermore, using the language of the standard does not guarantee alignment. No matter their wording, items must reflect the true intent of the standard in order to align properly.

Depth of Knowledge

Depth of knowledge (DOK) is a system developed by Dr. Norman L. Webb[1] for categorizing the cognitive demands required to produce an acceptable response. These categories of elements reflect the different levels of cognitive expectation, or depth of knowledge, that are required in order for the student to complete the task. As the cognitive demands increase, so does the DOK, ranging from a low of DOK 1 to a high of DOK 4.

A DOK 1 item would be the least demanding task and would require little more than recall or reproduction, such as solving a one-step mathematical problem using a well-defined algorithmic procedure. A DOK 4 item would demand extended thinking, involving analysis, synthesis, and planning, most likely over a period of time. As such, DOK 4 items are rare on assessments because test time is a critical factor; however, some extended-response tasks can reach this level of complexity.

[1]Webb, N. L. (1997). *Research monograph No. 6. Criteria for alignment of expectations and assessments in mathematics and science education.* Washington, DC: Council of Chief State School Officers.

High-quality standardized assessments have been moving away from the use of DOK 1 items. Given that test time and test "real estate" are at a premium, DOK 1 items do not provide enough psychometric discrimination to help distinguish between the students who have the knowledge and abilities the assessments are measuring and those who do not. In addition, DOK 1 items rarely achieve the level of rigor that is desirable in high-quality assessments. Therefore, most high-quality assessment items are designated as DOK 2 or DOK 3.

Mathematics DOK 2 items require students to make some decisions on how to solve a problem. Almost all such items require at least two steps to solve. Items at this level may require the student to compare, interpret, or classify information. For example, a problem might ask a student to compare two objects by interpreting information from a simple graph. Other examples could include identifying nontrivial patterns from a collection of data or carrying out a two-step computation in which the steps are not part of the same algorithm. Short constructed-response items that require explanations or simple interpretations are good examples of DOK 2 items, and many rigorous multiple-choice items are rated as DOK 2.

DOK 3 items requires students to reason, analyze, and connect concepts in a way that exceeds how these skills are used at the DOK 2 level. Items at the DOK 3 level may require students to make conjectures based on given or calculated values or to explain the mathematical reasoning that underlies a result. Most prompts that require a substantial written response are classified as DOK 3, in part because generating and presenting an extended response generally requires higher-level thinking skills than selecting from among several provided answer choices. However, some multiple-choice items that engage with complex or subtle ideas could also be designated as DOK 3.

It is important to note that DOK is a measure of cognitive complexity, not a measure of difficulty. In the realm of assessment, difficulty is determined by the percentage of students who correctly answer an item. It is possible for a DOK 2 item to be very difficult and for a DOK 3 item to be relatively easy.

Distractors

Incorrect answers, also known as distractors, are essential to the quality and rigor of an item. Weak distractors can be easily eliminated by test takers and do not provide useful information about misconceptions that a student may be struggling with. Strong distractors raise the rigor of an item by requiring the student to consider the merits of each answer choice, and can provide insight into student misconceptions.

Weak distractors seem implausible and may have little or no connection to the information given in the item stem, or may be clearly contradictory to the stem information. A distractor is weak if it can be automatically eliminated by a test taker who has not mastered the skill or concept that is being assessed.

Strong distractors may seem plausible and reflect likely misconceptions or common errors made when carrying out computations. A strong distractor might seem appealing to a student who has only partially mastered a skill or concept, or who has a minimal understanding of how to determine a necessary computation or how to carry out a necessary computation.

If it is not possible to write three strong distractors for a single-select multiple-choice item, the targeted material may not be rich enough to yield a successful item.

Distractors should not be potentially keyable or deliberately misleading. Care must be taken to distinguish between strong distractors and weak correct answers.

Construct-Irrelevant Sources of Challenge

While it is essential for items to be rigorous, it is important to differentiate between items that are complex and items that are merely complicated in their design.

In a well-crafted item, the source of the challenge should be essential to the construct that is being measured. In a poorly crafted item, construct-irrelevant sources of challenge may prevent the student from demonstrating the knowledge, skill, or ability that is being measured.

Construct-irrelevant sources of challenge often include:

High reading load: Items should be clear and concise. Remove unnecessary verbiage from item stems, prompts, and answer choices, and ensure that all vocabulary and mathematical terminology are on or (ideally) below grade level and are accessible. Students should be able to decode items with little effort. (See the grade K item exemplar on page 22.)

Idiomatic or figurative language: Construct-irrelevant instances of idiomatic or figurative language are particularly problematic for English language learners. Such language should be

replaced with wording that is more literal. In addition to enhancing accessibility, this practice ensures that the item is measuring what it is intended to measure.

The boldfaced following words or phrases are examples of potentially problematic language:

After the event, how many boxes will be **left over**? (It is better to say "remaining" to avoid two statements of direction.)

Sam **is running for** class president. (It is better to say "wants to be elected" to avoid a misinterpretation as the literal action of running.)

Unclear item directions: Providing clear directions is essential, particularly for multiple-select and technology-enhanced items. In multiple-select items, consider including in the stem the number of correct answers that a student must select. In technology-enhanced items, do not assume that students will intuitively know which steps to follow to complete the task. Provide directions that someone with little knowledge of the item format or the technology interface could successfully follow.

Assumptions of prior knowledge: Ensure that items do not require or assume knowledge that a student might not possess. For example, an application problem that requires the use of a scientific formula should state the formula and define the formula variables and constants. Care must also be taken when choosing contexts, particularly for items involving geometric figures. For example, rather than simply stating that a carton has a base area of 5 m^2 and a volume of 65 m^3, present a diagram of the carton with the given measurements, so that the student understands that the carton is a right rectangular prism.

Accessible Language

These additional best practices can help students from all populations access the material on an assessment.

Use familiar contexts. Consider whether students will be familiar with the key words and concepts that are used in an item. For example, in a real-world application requiring use of a familiar concept or formula, students should not be distracted or confused by unfamiliar vocabulary or abstract concepts tied to the context. Replace potentially unfamiliar, lower-frequency lexical items with more-familiar, higher-frequency lexical items. Additionally, care must be taken to ensure that the context is non-offensive and free of bias.

Use the present tense. Most items should be written in the present tense. Even items with contexts that deal with future events can be written in the present tense: e.g., "Determine how much money the student earns at the end of one week." rather than "Determine how much money the student will have earned at the end of one week."

Use the active voice. Passive verb constructions are more difficult for students, especially English language learners, to process and remember. Other languages use the passive voice much less frequently and/or in very limited situations, so English language learners may not have had much exposure to passive constructions.

Use closed stems. Closed stems are grammatically complete and are therefore easier for English language learners to interpret. It may not be immediately apparent to a student that the student is supposed to supply the ending to a grammatically incomplete open stem.

Use a positive frame of reference. Avoid stems that use negatives (e.g., "Which value is **not** a solution of the inequality?"). Students may not register the negative construction and may therefore respond incorrectly. Negative constructions are more difficult to process, and may introduce unnecessary confusion. In addition, using negative answer choices (e.g., "$x \neq 4$") with this example stem would create a double negative.

While it is generally preferable to use a positive frame of reference in all parts of the item, there may be instances in which it is desirable to use a negative construction in the answer choices. In these cases, ensure that the answer choices are balanced in this regard (e.g., two answer choices have negative constructions and two do not).

Ensure that pronouns and referents are clear. To minimize confusion about pronoun referents and to lessen the need for gender specification, use gender-neutral terms, such as "the student" or "a scientist," to refer to individuals in an item context.

Cueing and Clanging

Cueing and clanging are related issues that can lead students to select or eliminate an answer choice for construct-irrelevant reasons. It is important to craft items so that cueing is avoided entirely and clanging is minimized. In addition, items that are administered together must be reviewed carefully for cross-item instances of cueing and clanging.

Cueing occurs when one item gives away the answer to another item, or when a certain element within an item reveals the answer to that item.

A subtler form of cueing occurs when the wording or construction of the item is likely to make a certain response seem especially attractive to a test-taker—in other words, when the student may select the right answer for the wrong reason.

It is also possible for students to be miscued, or led to select an incorrect answer, for construct-irrelevant reasons.

Clanging is the repetition of a number, word, or phrase in a way that is likely to be distracting to a test taker (e.g., included in some distractors but not others).

Clanging can lead to cueing. For example, the use of a key number, word, or phrase in the stem and in one answer choice might lead the student to select that answer choice.

Item Development Checklist and Guiding Questions

Item Development Checklist

Item Set

- ✓ Cueing is avoided across the item set.
- ✓ A range of standards is assessed across the item set.
- ✓ A range of complexity and difficulty is assessed across the item set.

Individual Items

- ✓ Item stem is closed.
- ✓ The item aligns to the standard that it is intended to measure.
- ✓ The item assesses grade-appropriate content.
- ✓ The levels of complexity and difficulty are grade-appropriate.
- ✓ The stem is concisely worded and uses grade-appropriate vocabulary.
- ✓ No misleading wording, context, or extraneous numbers are used.
- ✓ The student's task is made clear in the stem. Scaffolding is used when appropriate.
- ✓ Graphics, if used, are clear and accurate.
- ✓ Clanging is avoided within the item.
- ✓ The item context, if used, is non-offensive and free of bias.

Answer Choices

- ✓ There is only one correct answer (except in multiple-select items).
- ✓ Options are parallel and balanced, and outliers (e.g., use of a value that is significantly different from those used in the other answer choices) are avoided.
- ✓ Options are grammatically and syntactically compatible with the stem.

- ✓ Options are ordered in accordance with the type(s) of information that they include.
- ✓ Each option is plausible, and each distractor supports a rationale explaining the mistake that the student most likely made in choosing that option.

Technology-Enhanced Items

- ✓ The use of technology is justified (the item allows the student to respond in a way that is not possible or is not efficient via a traditional multiple-choice item).
- ✓ The technological aspects of the item do not introduce unnecessary demands on students.
- ✓ The standard that the item assesses lends itself well to the use of the format.

Item Development Guiding Questions

- ✓ Does the item require the student to **draw on information given in the stem** in order to respond correctly?
- ✓ Is the item **appropriately complex** for the grade level?
- ✓ Does the item require the student to discern among **strong answer choices** in order to arrive at the correct answer?
- ✓ Is the item clearly and purposefully **aligned to the standard(s)**?
- ✓ Does the item **mirror the language of the standard(s)**, if feasible?

Item Exemplars: Grades K-5

Grade K

ITEM FORMAT: Single-Select Multiple Choice (SSMC)		
PRIMARY STANDARD: Louisiana State Standard K.OA.A.4	For any number from 1 to 9, find the number that makes 10 when added to the given number, e.g., by using objects or drawings, and record the answer with a drawing or equation.	
KEY: B	DOK: 2	CALCULATOR: No

ITEM:

Kate wants to pick 10 flowers. She has picked 6 flowers so far. How many **more** flowers does she need to pick?

A. 3
B. 4
C. 6
D. 7

Rationales:
A. used number that, when added to itself, equals the number of flowers picked so far
B. correct
C. used number of flowers picked so far
D. used next sequential number after 6

COMMENTS AND NOTES:

Mechanics:

- The answer choices in the item are balanced: two are even and two are odd.
- All four choices are parallel: all are one-digit numbers.
- Following best practice, the choices are arranged in ascending order.
- The language in the stem is simplified; short sentences present relevant information in pieces, making the item more appropriate for the grade level. There is no extraneous context or reading load.

Grade K (continued)

Content:

- The item strongly aligns to the standard.
- The item contains only one correct answer and has plausible/logical distractors.
- The context is simple and uses a common object that should be familiar to all students.
- The item's complexity is medium; the student must use context clues and the given aid (i.e., the emphasis word) to determine the correct operation required to solve, and then to perform the operation correctly.
- The item is designated as "Calculator: No" because students at this grade level are learning number sense and number facts, and a computational aid would hinder their ability to show mastery of the content.

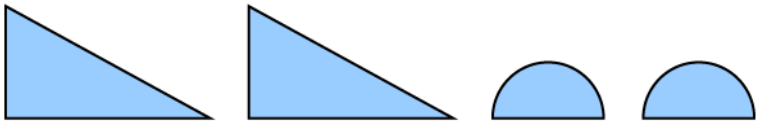
Grade 1

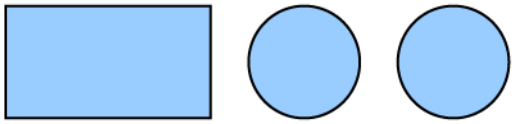
ITEM FORMAT: Single-Select Multiple Choice (SSMC)		
PRIMARY STANDARD: Oklahoma State Standard 1.GM.1.2	Compose and decompose larger shapes using smaller two-dimensional shapes.	
KEY: A	DOK: 2	CALCULATOR: Neutral

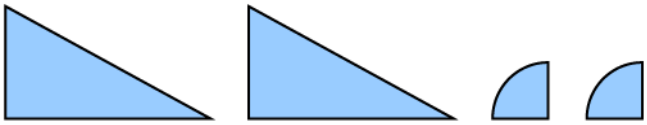
ITEM:

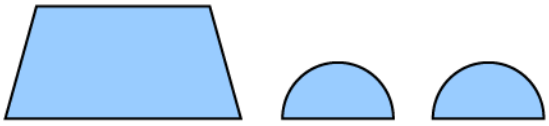
Some shapes are used to make the figure shown below. The shapes do not overlap.

Which group shows the shapes that could be used to make the figure?

A. 

B. 

C. 

D. 

Rationales:

A. correct
B. selected option with central rectangle, but circles would overlap the rectangle
C. selected option with triangles that form a central rectangle, but two additional quarter-circles would be needed to complete the figure
D. selected option with semicircles that form the outer portions of the figure, but central shape must be a rectangle

Grade 1 (continued)

Comments and Notes:

Mechanics:

- The stem graphic is introduced, and students are told information that is relevant to the task that they must complete.
- The answer choices in the item are balanced: two show 4 shapes and two show 3 shapes.
- The choices are arranged by length, in an alternating long/short order.
- The language in the stem is simple and direct; there is no extraneous context or reading load.

Content:

- The item strongly aligns to the standard.
- The item contains only one correct answer and has plausible/logical distractors. The second sentence in the introduction is critical, as choice B could also be considered correct if non-overlapping shapes are not defined as a criterion.
- The item's complexity is medium; the student must evaluate the shapes in each answer choice and use spatial reasoning to determine whether those shapes could be rearranged and placed side-by-side to compose the given figure.
- The item is designated as "Calculator: Neutral" because using a computational aid would neither help nor hinder a student's ability to show mastery of the content.

Grade 2

ITEM FORMAT: Multiple-Select Multiple Choice (MSMC)		
PRIMARY STANDARD: CCSS 2.NBT.4	Compare two three-digit numbers based on meanings of the hundreds, tens, and ones digits, using >, =, and < symbols to record the results of comparisons.	
KEY: A, B	DOK: 1	CALCULATOR: Neutral

ITEM:

One digit is missing in the number comparison shown below.

$$425 > 4 \square 5$$

Which of these could be the missing digit? Select **two** correct answers.

A. 0
B. 1
C. 2
D. 3
E. 4
F. 5

Rationales:
A. correct
B. correct
C. selected the digit that would make numbers equal
D. reversed direction of inequality symbol
E. reversed direction of inequality symbol; selected hundreds digit
F. reversed direction of inequality symbol; selected ones digit

COMMENTS AND NOTES:

Mechanics:

- The number comparison is introduced, and students are told information that is relevant to the task that they must complete, including how many answers to select.
- The answer choices in the item are balanced: three are even and three are odd.
- All six choices are parallel: all are one-digit numbers.
- Following best practice, the choices are arranged in ascending order.
- The language in the stem is simple and direct; there is no extraneous context or reading load.

Grade 2 (continued)

Content:

- The item strongly aligns to the standard.
- The item contains more than one correct answer and has plausible/logical distractors.
- The item's complexity is low; the cognitive demand placed on the student, who must perform a routine place-value comparison of two numbers, is low, but consideration that the number comparison could be true for multiple digits increases the difficulty of the task.
- The item is designated as "Calculator: Neutral" because using a computational aid would neither help nor hinder a student's ability to show mastery of the content.

Grade 3

ITEM FORMAT: Single-Select Multiple Choice (SSMC)		
PRIMARY STANDARD: CCSS 3.G.1	Understand that shapes in different categories (e.g., rhombuses, rectangles, and others) may share attributes (e.g., having four sides), and that the shared attributes can define a larger category (e.g., quadrilaterals). Recognize rhombuses, rectangles, and squares as examples of quadrilaterals, and draw examples of quadrilaterals that do not belong to any of these subcategories.	
KEY: C	DOK: 1	CALCULATOR: Neutral

ITEM:

Which shape is a quadrilateral?

A. a hexagon
B. a pentagon
C. a rhombus
D. a triangle

Rationales:
A. selected shape with 6 sides
B. selected shape with 5 sides
C. correct
D. selected shape with 3 sides

COMMENTS AND NOTES:

Mechanics:

- The language in the stem is simple and direct; there is no extraneous context or reading load.
- Following best practice, the choices are arranged in alphabetical order.

Content:

- The item strongly aligns to the second component of the standard.
- The item contains only one correct answer.
- The item's complexity is low; the student must identify the type of shape that has four sides and is therefore classified as a quadrilateral.
- The item is designated as "Calculator: Neutral" because using a computational aid would neither help nor hinder a student's ability to show mastery of the content.

Grade 3 (continued)

Alternative Item:

Item Format: TEI—Multiple-Select Grid		
Primary Standard: CCSS 3.G.1	Understand that shapes in different categories (e.g., rhombuses, rectangles, and others) may share attributes (e.g., having four sides), and that the shared attributes can define a larger category (e.g., quadrilaterals). Recognize rhombuses, rectangles, and squares as examples of quadrilaterals, and draw examples of quadrilaterals that do not belong to any of these subcategories.	
Key: [see art]	DOK: 2	Calculator: Neutral

Item:

For each shape shown in the table below, show whether the shape is always a parallelogram, always a rectangle, always both a parallelogram and a rectangle, or never a parallelogram or a rectangle.

Select one cell in **each** row of the table.

Shape	Is Always a Parallelogram	Is Always a Rectangle	Is Always Both a Parallelogram and a Rectangle	Is Never a Parallelogram or a Rectangle
rhombus	○	○	○	○
square	○	○	○	○
trapezoid	○	○	○	○

Comments and Notes:

Key:

Shape	Is Always a Parallelogram	Is Always a Rectangle	Is Always Both a Parallelogram and a Rectangle	Is Never a Parallelogram or a Rectangle
rhombus	●	○	○	○
square	○	○	●	○
trapezoid	○	○	○	●

Grade 3 (continued)

Mechanics:

- The table is introduced, and students are told information that is relevant to the task that they must complete.

Content:

- The item strongly aligns to the first component of the standard.
- The item contains only one correct answer per row.
- The item's complexity is medium; the complexity is increased because the student must classify each shape by considering their attributes, and must consider that a shape may be classified as fitting into more than one category. This task is made slightly more difficult without pictures, as the student must conceptualize all possible forms of each shape.
- The item is designated as "Calculator: Neutral" because using a computational aid would neither help nor hinder a student's ability to show mastery of the content.

Grade 4

ITEM FORMAT: Constructed-Response Short-Answer		
PRIMARY STANDARD: CCSS 4.OA.3	Solve multistep word problems posed with whole numbers and having whole-number answers using the four operations, including problems in which remainders must be interpreted. Represent these problems using equations with a letter standing for the unknown quantity. Assess the reasonableness of answers using mental computation and estimation strategies including rounding.	
KEY: Rubric	DOK: 3	CALCULATOR: No

ITEM:

John works at an apple orchard. He collects a total of 252 pounds of apples.

- He collects 86 pounds from one field.
- He collects 44 pounds from a second field.
- He collects the remainder of the apples from a third field.

John keeps 30 pounds of apples that he collects from the third field for himself and puts the remaining apples that he collects from the third field into boxes. Each box holds up to 9 pounds of apples.

What is the **fewest** number of boxes that John needs to hold the remaining apples from the third field? Explain how you arrived at your answer.

SAMPLE SCORING RUBRIC:

Score Points:
1 point: correct answer
1 point: valid explanation given

Sample Answer:
The fewest number of boxes that John needs to hold the remaining apples from the third field is 11.

To find the number of apples that John collects in the third field, I added 86 and 44, and then subtracted the result from 252.

$86 + 44 = 130$

$252 - 130 = 122$

Then, since John keeps 30 pounds of apples that he collects from the third field for himself, I subtracted.

$122 - 30 = 92$

Since each box holds up to 9 pounds of apples, I divided 92 by 9.
$92 \div 9 = 10R2$

There will be a total of 10 full boxes needed for 90 pounds of apples and 1 additional box needed to hold the remaining 2 pounds of apples.
$10 + 1 = 11$

[Accept other valid responses.]

Comments and Notes:

Mechanics:

- Information in the stem is shown in a bulleted list, which presents relevant information to students in pieces.
- The question is set apart on its own line to visually distinguish the task from the information to be used to complete the task.

Content:

- The item strongly aligns to the first and last components of the standard.
- The item's complexity is high; students must synthesize the information presented by using context clues and the given aid (i.e., emphasis word), determine and perform the operations needed to complete the task, and then explain their process and/or reasoning.
- The context uses a common object that should be familiar to all students.
- The item is designated as "Calculator: No" because students at this grade level are responsible for performing the operations with whole numbers correctly, but are not yet responsible for interpreting decimal remainders that a computational aid would likely produce.

Grade 5

ITEM FORMAT: Single-Select Multiple Choice (SSMC)		
PRIMARY STANDARD: CCSS 5.NF.4a	Apply and extend previous understandings of multiplication to multiply a fraction or whole number by a fraction. • Interpret the product $(a/b) \times q$ as a parts of a partition of q into b equal parts; equivalently, as the result of a sequence of operations $a \times q \div b$. *For example, use a visual fraction model to show (2/3) × 4 = 8/3, and create a story context for this equation. Do the same with (2/3) × (4/5) = 8/15. (In general, (a/b) × (c/d) = ac/bd.)*	
KEY: A	DOK: 2	CALCULATOR: No

ITEM:

Which statement describes the product of $\frac{4}{9}$ and $\frac{5}{8}$?

A. 4 groups of $\frac{5}{8}$ divided into 9 equal parts

B. 9 groups of $\frac{5}{8}$ divided into 4 equal parts

C. 4 groups of $\frac{5}{8}$ divided into 72 equal parts

D. 9 groups of $\frac{5}{8}$ divided into 20 equal parts

Rationales:
A. correct
B. reversed the functions of the numerator and the denominator
C. multiplied the denominators
D. reversed the functions of the numerator and the denominator and multiplied the numerators

Grade 5 (continued)

Comments and Notes:

Mechanics:

- The answer choices in the item are balanced: two begin with 4 and two begin with 9.
- The choices are arranged in pairs and ordered from least to greatest by the first digit that appears in each pair.
- The language in the stem is simple and direct; there is no extraneous context or reading load.

Content:

- The item strongly aligns to the standard.
- The item contains only one correct answer and has plausible/logical distractors.
- The item's complexity is medium; by interpreting the product, rather than performing a routine computation of the product, students are demonstrating a deeper understanding of the meaning of multiplying a fraction by a fraction.
- The item is designated as "Calculator: No" because students at this grade level are learning fractional number sense, and also to keep the focus on the conceptual interpretation of the product.

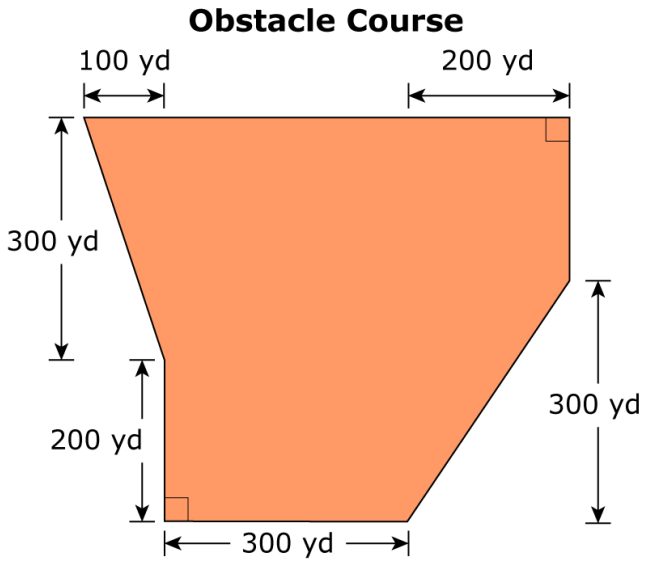
Item Exemplars: Grades 6-HS

Grade 6

ITEM FORMAT: Constructed-Response Short Answer		
PRIMARY STANDARD: CCSS 6.G.1	Find the area of right triangles, other triangles, special quadrilaterals, and polygons by composing into rectangles or decomposing into triangles and other shapes; apply these techniques in the context of solving real-world and mathematical problems.	
SECONDARY STANDARD: CCSS 6.RP.3b	Use ratio and rate reasoning to solve real-world and mathematical problems, e.g., by reasoning about tables of equivalent ratios, tape diagrams, double number line diagrams, or equations. • Solve unit rate problems including those involving unit pricing and constant speed. *For example, if it took 7 hours to mow 4 lawns, then at that rate, how many lawns could be mowed in 35 hours? At what rate were lawns being mowed?*	
KEY: Rubric	DOK: 3	CALCULATOR: Yes

ITEM:

An event manager creates an obstacle course for an upcoming event. The dimensions of the outside edge of the obstacle course are shown below.



Grade 6 (continued)

- The event manager determines that about $\frac{4}{5}$ of the obstacle course will be used for obstacles.
- Every 100 square yards of the obstacle course that are **not** used for obstacles will hold about 400 people who can observe the event.

Determine the maximum number of people who can observe the event. Show your work or explain your response.

Sample Scoring Rubric:

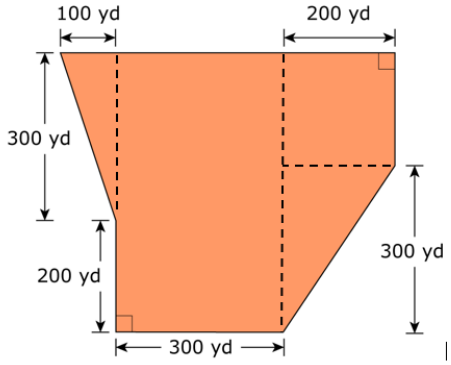
Score Points:

1 point: correct answer
1 point: valid explanation given

Sample Answers:

The maximum number of people who can observe the event is about 188,000.

To determine the area of the obstacle course, the course can be divided into rectangles and triangles, as shown below.



- The area of the triangle in the upper left is 15,000 yd², since $0.5(100 \bullet 300) = 15{,}000$.
- The area of the large rectangle is 150,000 yd², since $300(300 + 200) = 150{,}000$.
- The area of the square in the upper right is 40,000 yd², since $200(300 + 200 - 300) = 40{,}000$.
- The area of the triangle in the lower right is 30,000 yd², since $0.5(200 \bullet 300) = 30{,}000$.

The total area of the obstacle course is 235,000 yd², since
$15{,}000 + 150{,}000 + 40{,}000 + 30{,}000 = 235{,}000$.

Grade 6 (continued)

Because obstacles will cover $\frac{4}{5}$ of the course, $\frac{1}{5}$ can be used for people to observe, since $1-\frac{4}{5}=\frac{1}{5}$.

Because $235{,}000\left(\frac{1}{5}\right)=47{,}000$, there are 47,000 yd² that can be used for people to observe, and because $47{,}000\div 100=470$, there are 470 sections of 100 yd². If each of those sections can hold about 400 people, then about 188,000 people can observe the event, since $\frac{400}{100}=\frac{188{,}000}{47{,}000}$.

[Accept other valid responses.]

Comments and Notes:

Mechanics:

- The stem graphic is introduced, and students are told information that is relevant to the task that they must complete.
- Information in the stem is shown in a bulleted list, which presents relevant information to students in pieces.
- The prompt is set apart on its own line to visually distinguish the task from the information to be used to complete the task.

Content:

- The item strongly aligns to more than one standard. The primary standard is considered primary due to the fact that ratio/rate reasoning cannot be used to answer the item successfully without an understanding of how to decompose a polygon and determine the area.
- The item's complexity is high; there are numerous ways to decompose the given polygon, and the student must identify a way in which the given dimensions will allow the area to be determined. Students must also be able to explain their process and/or reasoning. Additionally, there is no scaffolding for the student to use to determine either the correct operations to perform or the correct order in which to perform them, which contributes to the high complexity rating.
- The item is designated as "Calculator: Yes" because the computations themselves are not the source of challenge for the item. Students should be able to use a calculator to efficiently compute an answer, allowing more time to be spent on determining the correct sequence of steps and on explaining the validity of that sequence.

Grade 7

ITEM FORMAT: Single-Select Multiple Choice (SSMC)		
PRIMARY STANDARD: CCSS 7.SP.8c	Find probabilities of compound events using organized lists, tables, tree diagrams, and simulation. • Design and use a simulation to generate frequencies for compound events. *For example, use random digits as a simulation tool to approximate the answer to the question: If 40% of donors have type A blood, what is the probability that it will take at least 4 donors to find one with type A blood?*	
KEY: B	DOK: 2	CALCULATOR: Yes

ITEM:

Approximately 20% of Americans play a musical instrument. A student designs an accurate simulation to determine how many people would need to be randomly selected before 5 people who play a musical instrument are found. The student creates a spinner with 10 equal-sized sections and will record the results of spinning the arrow on the spinner. How many of the equal-sized sections should be used to represent Americans who play a musical instrument?

A. 1
B. 2
C. 8
D. 9

Rationales:
A. determined that 1 in 5 Americans play a musical instrument
B. correct
C. determined the number of sections that represent Americans who do not play a musical instrument
D. determined the number of sections that represent Americans who do not play a musical instrument when 1 section represents Americans who do play a musical instrument

COMMENTS AND NOTES:

Mechanics:

- The answer choices in the item are balanced: two are even and two are odd.
- All four choices are parallel: all are one-digit numbers.
- Following best practice, the choices are arranged in ascending order.

Grade 7 (continued)

Content:

- The item strongly aligns to the standard.
- The item contains only one correct answer and has plausible/logical distractors.
- The item's complexity is medium; the student must interpret and analyze the information given and determine how to use the information within the bounds of the defined simulation to generate a specified probability.
- The item is designated as "Calculator: Yes" because the computations themselves are not the source of challenge for the item. Students should be able to use a calculator to efficiently compute an answer, allowing more time to be spent on determining the correct simulation for a specific probability.

Grade 8

ITEM FORMAT: Single-Select Multiple Choice (SSMC)		
PRIMARY STANDARD: Pennsylvania Eligible Content M08.A-N.1.1.4	Use rational approximations of irrational numbers to compare and order irrational numbers.	
KEY: C	DOK: 2	CALCULATOR: No

ITEM:

The inequality below shows that the value of *w* is less than the irrational number $3\sqrt{6}$.

$$w < 3\sqrt{6}$$

Which inequality shows a possible value of *w* and an approximate value of $3\sqrt{6}$ that make the inequality true?

A. $4\pi < 7$
B. $7\pi < 18$
C. $\sqrt{20} < 7$
D. $\sqrt{27} < 18$

Rationales:

A. determined an approximate value of $3\sqrt{6}$, but reversed direction of inequality
B. multiplied 3 by 6 to determine an approximate value of $3\sqrt{6}$; reversed direction of inequality
C. correct
D. inequality is true, but multiplied 3 by 6 to determine an approximate value of $3\sqrt{6}$

COMMENTS AND NOTES:

Mechanics:

- The expression is introduced, and students are told information that is relevant to the task that they must complete.
- The expression is set apart on its own line to visually set it apart from the surrounding text.
- The answer choices in the item are balanced: two use 7 and two 18 as approximations of $3\sqrt{6}$.
- The choices are not arranged in ascending order in this case, as that could cue that choice A contains an expression with the least value. Instead, the choices are ordered by the value of the digits on the left side of each inequality (but not the value of the irrational numbers), so no choices can be eliminated due to their position.

Grade 8 (continued)

Content:

- The item strongly aligns to the standard.
- The item contains only one correct answer and has plausible/logical distractors.
- The item's complexity is medium; students must approximate not only the given irrational number but also the irrational numbers presented in the answer choices. They must consider whether the irrational numbers satisfy the parameters for values of *w*, as well as whether the inequalities are true.
- The item is designated as "Calculator: No" to keep the focus on the estimation of the values of irrational numbers and on number sense.

Grade 8 (continued)

Alternative Item:

Item Format: Single-Select Multiple Choice (SSMC)		
Primary Standard: Pennsylvania Eligible Content M08.A-N.1.1.4	Use rational approximations of irrational numbers to compare and order irrational numbers.	
Key: B	DOK: 3	Calculator: No

Item:

Three expressions are listed below. In each expression, *w* represents a number between 0 and 1.

$$\sqrt{w},\ \frac{\sqrt{w}}{4},\ \frac{1}{\sqrt{w}}$$

Which list shows the expressions in order from **least** value to **greatest** value when *w* represents any number between 0 and 1?

A. $\frac{\sqrt{w}}{4},\ \frac{1}{\sqrt{w}},\ \sqrt{w}$

B. $\frac{\sqrt{w}}{4},\ \sqrt{w},\ \frac{1}{\sqrt{w}}$

C. $\frac{1}{\sqrt{w}},\ \sqrt{w},\ \frac{\sqrt{w}}{4}$

D. $\frac{1}{\sqrt{w}},\ \frac{\sqrt{w}}{4},\ \sqrt{w}$

Rationales:

A. considered the values of expressions with a numerator/denominator to be less than the value of $\sqrt{w}$

B. correct

C. reasoned with numbers greater than 1; considered the value of $4\sqrt{w}$ instead of $\frac{\sqrt{w}}{4}$

D. reasoned with numbers greater than 1

Grade 8 (continued)

Comments and Notes:

Mechanics:

- The expressions are introduced, and students are told information that is relevant to the task that they must complete.
- The expressions are set apart on their own line to visually set them apart from the surrounding text.
- The answer choices in the item are balanced: two begin with $\frac{\sqrt{w}}{4}$ and two begin with $\frac{1}{\sqrt{w}}$.
- The choices are not arranged in ascending order in this case, as that could cue that choice A begins with the expression with the least value and that the choices are ordered from least to greatest. Instead, the order is random, so that no choices can be eliminated due to their position.

Content:

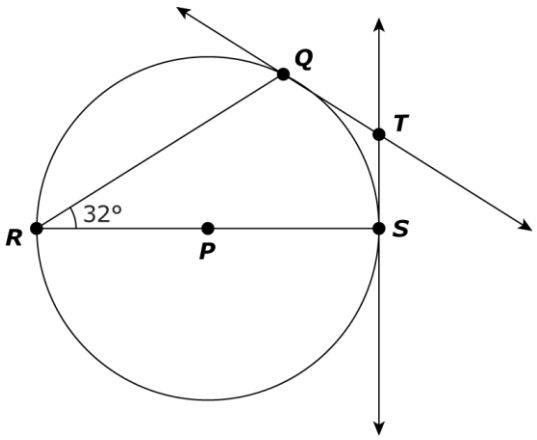
- The item strongly aligns to the standard.
- The item contains only one correct answer and has plausible/logical distractors.
- The item's complexity is high; as in the DOK 2 item, students must consider the variable *w* when comparing irrational numbers. However, in this item, they must also consider the rational value of *w* as a potentially irrational number when evaluating each expression, as well as considering whether their assumption for one specific rational value of *w* is true for the entire given range of values for *w*. They must make these considerations before being able to compare the values of the different expressions and place the expressions in order by value.
- The item is designated as "Calculator: No" to keep the focus on the estimation of the values of irrational numbers and on number sense.

Grade HS—Geometry

ITEM FORMAT: Single-Select Multiple Choice (SSMC)		
PRIMARY STANDARD: CCSS G-C.2	Identify and describe relationships among inscribed angles, radii, and chords. *Include the relationship between central, inscribed, and circumscribed angles; inscribed angles on a diameter are right angles; the radius of a circle is perpendicular to the tangent where the radius intersects the circle.*	
KEY: C	DOK: 2	CALCULATOR: Yes

ITEM:

Circle *P* and quadrilateral *QRST* are shown below.



- Line *ST* is tangent to circle *P* at point *S*.
- Line *QT* is tangent to circle *P* at point *Q*.
- The measure of angle *QRS* is 32°.

What is the measure of angle *QTS*?

A. 32°
B. 64°
C. 116°
D. 148°

Rationales:
A. reasoned that the measure of angle *QTS* equals the measure of angle *QRS*
B. reasoned that the measure of angle *QTS* equals the measure of the central angle
C. correct
D. reasoned that the measure of angle *QTS* equals 180° minus the measure of angle *QRS*

Comments and Notes:

Mechanics:

- The stem graphic is introduced, and students are told information that is relevant to the task that they must complete.
- Information in the stem is shown in a bulleted list, which presents relevant information to students in pieces.
- The stem graphic provides clarity about the information and relationships shown in the bulleted list.
- The answer choices in the item are balanced: two are 2 digits and two are 3 digits.
- Following best practice, the choices are arranged in ascending order.

Content:

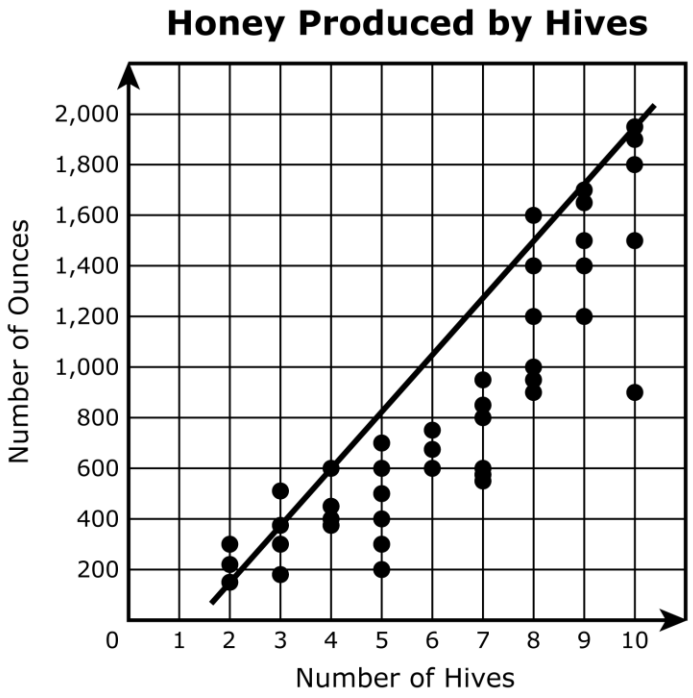
- The item strongly aligns to the standard.
- The item contains only one correct answer and has plausible/logical distractors.
- The item's complexity is medium; students must decode how the information in the bulleted list will help them solve the task, and must know which circle theorems to apply.
- The item is designated as "Calculator: Yes" because angle relationships, not computation, are the focus of the item. Students should be able to use a calculator to efficiently compute an answer, allowing more time to be spent on determining the correct relationships to use to solve the answer.

Grade HS—Statistics

ITEM FORMAT: Constructed Response (CR)		
PRIMARY STANDARD: Texas Essential Knowledge and Skills 111.32.b.2.D	Collect and organize data, make and interpret scatterplots (including recognizing positive, negative, or no correlation for data approximating linear situations), and model, predict, and make decisions and critical judgments in problem situations.	
SECONDARY STANDARD: Texas Essential Knowledge and Skills 111.47.c.7.F	Identify and interpret the reasonableness of attributes of lines of best fit within the context, including slope and *y*-intercept.	
KEY: Rubric	DOK: 3	CALCULATOR: Yes

ITEM:

The scatter plot below uses points to display the relationship between the number of hives that a beekeeper tended in a week and the total number of ounces of honey that those hives produced.



A. State and interpret the association between the number of hives that a beekeeper tended in a week and the total number of ounces of honey that those hives produced.

B. Explain why the line segment shown on the graph does or does not represent the line of best fit for the relationship displayed. As part of your explanation, make and justify a prediction of the total number of ounces of honey that the beekeeper should expect 14 hives to produce when the beekeeper tends to those hives in a week.

A different scatter plot is created to display the relationship between the number of ounces of honey that a hive produced and the profit, in dollars, from the sale of the honey. A point on the scatter plot lies above the line of best fit.

C. In context, interpret the meaning of a point that lies above the line of best fit.

Sample Scoring Rubric:

Score	Description
4	Student scores 4 points.
3	Student scores 3–3.5 points.
2	Student scores 2–2.5 points.
1	Student scores 0.5–1.5 points.
0	Student's response is incorrect or provides insufficient evidence of appropriate skills or knowledge to successfully accomplish the task.
Blank	No student response.

Score Points

A.	score 1 point **OR** score 0.5 point	correct answer with correct and complete explanation correct answer with incomplete explanation **OR** vague explanation only
B.	score 2 points **OR** score 1 point **OR** score 0.5 point	correct answer and correct prediction with complete explanations correct answer and correct prediction with incomplete or no explanation correct answer only or correct prediction only with incomplete or no explanation **OR** vague explanation only
C.	score 1 point **OR** score 0.5 point	correct and complete explanation incomplete or vague explanation

Sample Answers

A. As the number of hives tended in a week increases, the total number of pounds of honey produced increases. Therefore, the association is positive, and the line of best fit will have a positive slope.

[Accept other valid responses.]

B. The line segment shown on the graph does not represent the line of best fit. The line of best fit most closely approximates the data. As most of the points on the scatter plot are below the line segment, the line segment does not represent the data well.

It appears that the line of best fit would come closer to passing through the points (5, 600) and (10, 1500). Using these two points, the slope of the line would be about 180, since $\frac{(1500 - 600)}{(10 - 5)} = \frac{900}{5} = 180$. Then, the total number of ounces of honey produced when the beekeeper tends to 14 hives would be about 2,200, since $1500 + 180(14 - 10) = 2200$.

[Accept other valid responses.]

C. A point that lies above the line of best fit would indicate a greater profit than expected for the given number of ounces of honey produced.

[Accept other valid responses.]

Comments and Notes:

Mechanics:

- The stem graphic is introduced, and students are told information that is relevant to the task that they must complete.
- Although the item is worth 4 points, there are deliberately fewer than four parts/tasks, in order to leverage the format of the item and to elicit more in-depth responses than could be achieved by assessing four MC items.

Grade HS—Statistics (continued)

Content:

- The item strongly aligns to the standard.
- The item's complexity is high; the introduction states that there is a relationship shown in the scatter plot, but the student must analyze the scatter plot to determine what that relationship is, be able to describe it, and be able to justify and support the conclusions drawn from the data. The student must also explain phenomena in terms of concepts by relating specific data to the context in which they are set.
- The item is designated as "Calculator: Yes" because using a computational aid would not hinder a student's ability to show mastery of the content and could facilitate the prediction asked for in part B.