

# Smarter Balanced

## Mathematics Performance Task Specifications

Revised March 2018



## TABLE OF CONTENTS

### DESIGN SPECIFICATIONS

A. Key Features of Mathematics Performance Tasks.....	2
B. Gatekeeper Criteria for Performance Assessments.....	2

### TASK DEVELOPMENT AND SCORING SPECIFICATIONS

C. Task Type Considerations .....	5
D. Blueprint .....	5
1. Stimuli Presentation.....	5
2. Sets of Items.....	5
3. Ramping and Decision Making .....	5
4. Independent vs. Interdependent Items .....	6
5. Alignment to Claims .....	7

## DESIGN SPECIFICATIONS

A Performance Task (PT) is an item type designed to provide students with an opportunity to demonstrate their ability to apply their knowledge and higher-order thinking skills to explore and analyze a complex, real-world scenario. A mathematics performance task elicits evidence of students' ability to "bring it all together" to develop a solution plan to the central challenge of the task.

The first section of this document defines A) key features, B) gatekeeper criteria, and C) criteria descriptors to guide the development and review of performance tasks for the Smarter Balanced Assessment Consortium. Task types (equivalent to genres in ELA) in mathematics include *Plan and Design*, *Analysis and Theory* and *Evaluate and Recommend* and hybrids of these types. Technical considerations for each task type are found in sections D–K of this document. The design specifications in this document are for performance tasks in mathematics and work in tandem with other Consortium-approved specifications, including mathematics content specifications, grade-level areas of emphases, universal design, general accessibility and bias guidelines, metadata requirements, and style guidelines for mathematics text, equations, and artwork.

### A. Key Features of Mathematics Performance Tasks

(adapted from *Smarter Balanced Performance Task Specifications* document)

Performance tasks should:

- integrate knowledge and skills across multiple Claims and Targets—a key component of college and career readiness.
- measure capacities such as depth of understanding, research skills, and/or complex analysis with relevant evidence.
- require student-initiated planning, management of information/data and ideas, and/or interaction with other materials.
- reflect a grade level, developmentally appropriate real-world problem. Tasks elements (data sets, values provided, etc.) are not restricted to those actualized in the real world, but these elements should be realistic.
- allow for multiple approaches.
- represent content that is relevant and meaningful to students.
- allow for demonstration of twenty-first-century skills, such as critically analyzing and synthesizing information presented in a variety of formats, media, etc.
- require scoring that focuses on the essence of the Claim(s) and Targets for which the task was written.
- be feasible for the school/classroom environment.
- allow for calculator use in grades 6, 7, 8 and 11.

### B. Gatekeeper Criteria for Performance Assessments

A set of cross-cutting criteria have been developed to guide the development and review of Performance Tasks. The term *gatekeeper* indicates these as essential components of a quality PT. Performance Tasks that do not meet these criteria would not be accepted into the PT item pool. The gatekeeper criteria are listed below.

<p>Aligned with Claims and Standards</p>	<p>PTs should go to the heart of the key Claims and Common Core State Standards for Mathematics. In particular, they should elicit evidence of Claims 2, 3, &amp; 4:</p> <ul style="list-style-type: none"> <li>○ Students can solve a range of complex, well-posed problems in pure<sup>1</sup> and applied mathematics, making productive use of knowledge and problem solving strategies. (Claim 2)</li> <li>○ Students can clearly and precisely construct viable arguments to support their own reasoning and to critique the reasoning of others. (Claim 3)</li> <li>○ Students can analyze complex, real-world scenarios and can construct and use mathematical models to interpret and solve problems. (Claim 4)</li> </ul> <p>In addition, they should elicit evidence of student engagement in the Common Core Mathematical Practices.</p>
<p><b>Developmentally Appropriate</b></p>	<p>PT topics, tasks, and scoring should be appropriate for the age and developmental experience base of the students.</p>
<p><b>Engaging</b></p>	<p>Topics should be authentic and realistic, engaging students in solving a problem or making a decision they would find relevant.</p>
<p><b>Accessible</b></p>	<p>Topics and tasks should minimize sources of bias, allow for multiple pathways, and provide appropriate scaffolds or supports while keeping in mind that sources and response types need to allow access for students with different English language proficiency and students with disabilities.</p>
<p><b>Purposeful and Coherent</b></p>	<p>Tasks should have an authentic purpose, and all task components should be connected to achieving that goal.</p>

---

<sup>1</sup> Note: PTs in particular aim to assess problem solving in applied mathematics.

## TASK DEVELOPMENT AND SCORING SPECIFICATIONS

Task writers should ensure that the mathematics of the task is correct, and should use precise mathematical language.

### C. Task Type Considerations

Plan and Design:

- For plan-and-design tasks, design constraints and parameters should be clearly labeled and explained in the prompt.

Evaluate and Recommend:

- For evaluate-and-recommend tasks, data sets should be drawn from authentic data sets. If they are mocked up, they should conform to reasonable estimates.
- The audience and format of the recommendation should be clearly stated (e.g., “Write a letter to your school principal with your recommendation.”)

Analysis and Theory:

- For analysis-and-theory tasks, data sets should include an appropriate amount of data given the expectation that students will develop a theory about relationships within the problem.

These tasks provide a natural opportunity to engage students in the refinement of their theories, as additional data can be provided which suggests another theory.

### D. Blueprint

#### 1. Stimuli Presentation

When presenting stimuli, the following guidelines apply:

- There should be a reference (using bold font as an indicator) that connects items to the specific stimulus resource required for a response. The name of the resource should be bold in both the stem and stimulus (e.g., Use **Table 1** to answer this question).
- The number of resources (tables/graphics) within the stimulus of a PT should be limited for grades 3–5, as follows:
  - Grade 3 – two or fewer
  - Grades 4–5 – three or fewer

#### 2. Sets of Items

Each Grades 3 - 7 Performance Task should consist of a set of four to six items and each Grades 8 & 11 Performance Task should consist of a set of three to five<sup>2</sup> items (total points not critical). Each item may be worth 0-4 points, with guidelines for awarding 0, 1, 2, 3, or 4 points, but no half points. Scoring guides should allow for partial credit.

#### 3. Ramping and Decision Making

Grades 3–7 PTs should exhibit appropriate ramping across items, while higher grades will

---

<sup>2</sup> At grades 8 & 11, a hand scored proof or justification extended response item aligned to Claim 3 will be included in the examination.

exhibit less ramping with each increasing grade. A ramped sequence of items should provide all students access to a task, while maintaining the challenge of the stated goal of the task for higher performers. A ramped sequence of items will begin with items that encourage entry into the task but are still directly related to the stated goal of the task. All PT items should focus on problem solving, reasoning, and modeling as opposed to procedural and computational skills.

While ramping is an appropriate part of building toward autonomous un-fragmented chains of reasoning in Grades 3 – 7, by Grades 8 & 11, tasks should utilize less ramping, be less closely guided, and require more autonomous decision-making.

The number of items in a PT should correspondingly decrease as the grade level increases, indicated by the aforementioned decrease from 4 – 6 items per PT in Grades 3-7 to 3-5 item per PT in Grades 8 & 11.

#### 4. Independent vs. Interdependent Items

- Guidelines for the first 1-2 items:
  - a. The first 1-2 items must be independent (i.e., not needed to score subsequent parts of the task), but still directly related to the stated goal of the task.
  - b. The first 1-2 items should provide entry into the task. Here, “entry” means having low to low-medium difficulty and encourage students to make sense of the stated goal of the task.
- Guidelines for the remaining items:
  - a. The remaining items in all performance tasks may be hand-scorable.
  - b. The remaining items cannot depend on the independent item(s) mentioned above.
  - c. The remaining items should be cohesive and may be interdependent with each other.
  - d. Rubrics for the interdependent items should explicitly prevent students from being penalized multiple times if a mistake made on a preceding item is correctly carried through subsequent items.
  - e. At grades 3–7, a minimum of two and a maximum of four items should be hand scored.
  - f. At grade 8 & 11, a minimum of one and a maximum of four items should be hand scored.

**5. Alignment to Claims**

Performance Task items should contribute scores to each claim according to the following distribution. Again, there should be a total of 4-6 items in grades 3-7 PTs, and 3-5 items in grades 8 & 11 PTs.

Claim/Score Reporting Category	PT Items
1. Concepts and Procedures	0
2. Problem Solving	1-2
3. Communicating Reasoning	1-2
4. Modeling and Data Analysis	1-3
Total	3-6 items