Smarter Balanced Assessment Consortium:

Performance Task Specifications

Developed by Measured Progress/ETS Collaborative

April 16, 2012

Smarter Balanced Assessment Consortium: Performance Task Specifications

Role of Smarter Balanced Performance Tasks

Taken during the final 12 weeks of the school year, the Smarter Balanced summative assessments for accountability will have two components: a comprehensive end-of-year computer adaptive assessment and performance tasks. These assessments in English language arts (ELA) and mathematics will provide measures of students' achievement (proficiency in meeting grade-level standards), academic growth, and progress toward college and career readiness. The focus of both assessment components will be the claims and targets identified in the Smarter Balanced content specifications for ELA/literacy and mathematics, which serve as "bridge documents" between the Common Core State Standards and the Smarter Balanced summative assessments. Through the use of technology and innovative item and task formats, the Smarter Balanced assessments will exemplify "next generation assessments," significantly improving upon traditional, large-scale accountability tests in terms of authenticity, accessibility, and coverage of skills that are identified in college and career standards (e.g., mathematics practices, problem solving, speaking and listening, use of technology), as described in the Smarter Balanced content specifications.

The domain of performance assessment is quite broad, encompassing a range of non–selected-response tasks. A Smarter Balanced performance task involves significant interaction of students with stimulus materials and/or engagement in a problem solution, ultimately leading to an exhibition of the students' application of knowledge and skills, often in writing or spoken language. Stimuli include a variety of information forms (e.g., readings, video clips, data), as well as an assignment or problem situation. A key component of college and career readiness is the ability to integrate knowledge and skills across multiple content standards. Smarter Balanced will address this ability through performance tasks, because it cannot be adequately assessed with selected-response or constructed-response items.

Smarter Balanced Performance Task Specifications

General Guidelines for Development of Performance Tasks

The Smarter Balanced Performance Task Work Group has offered general guidance for the development of Smarter Balanced performance tasks by identifying the essential characteristics of tasks, listed below. These guidelines also serve as reminders of good instructional practices.

A performance task must:

- **Integrate knowledge and skills across multiple content standards or strands within a content area** (RTTT, 2010; Khattri and Sweet, 1996).
- **Measure capacities such as depth of understanding, research skills, complex analysis, and identification/providing of relevant evidence** (Darling-Hammond and Pecheone, 2010; RTTT, 2010; Wood, Darling-Hammond, Neill, and Roschewski, 2007; Ayala, Shavelson, Yin, and Schultz, 2002).
- **Require student-initiated planning, management of information and ideas, interaction with other materials** (RTTT, 2010; Black and Wiliam, 1998; Wood, Darling-Hammond, Neill, and Roschewski, 2007). In the reading and writing tasks, students have an opportunity to plan their responses and manage and interact with information/data gained through reading or listening to/viewing texts. In mathematics, students determine and employ strategies for solving problems and use a variety of mathematical tools and techniques in doing so.
- **Require production of extended responses, such as oral presentations, exhibitions, and other scorable products, including more extended written responses which might be revised and edited** (RTTT, 2010; Wiggins, 1989; Ayala, Shavelson, Yin, and Schultz, 2002).
- **Reflect a real-world task and/or scenario-based problem** (Darling-Hammond, 1997; Wiggins, 1998). Performance tasks should incorporate real-world, college- and career-related skills that require students to accomplish complex goals during multiple testing sessions. Tasks should be multi-stepped and allow for reflection and revision.
- **Allow for multiple approaches** (Reed, 1993; Eisner, 1999). Writing or speaking tasks should encourage or allow multiple approaches to developing and organizing ideas. For example, narrative writing might be used to support the presentation of an argument, while analysis and synthesis might be used to convey ideas in a narrative. In mathematics, problems presented in tasks should lend themselves to multiple solutions and/or solution strategies.
- **Represent content that is relevant and meaningful to students** (Henderson, Karr, and Kidwell, 1998). These attributes can be verified during the piloting stage of the development process.
- **Allow for demonstration of important knowledge and skills, including those that address 21st-century skills such as critically analyzing and synthesizing information presented in a variety of formats, media, etc.** (CCSSI, 2010; Darling-Hammond and Pecheone, 2010). Performance tasks are really evidence that a student has collected all the relevant information necessary across years to successfully engage in the current grade-level standards. Thus, these tasks will incorporate knowledge and skills of prior grades by necessity, even though the major focus is on the standards for the current grade level.

- **Allow for multiple points of view and interpretations** (Dana and Tippins, 1993; Eisner, 1999). In both ELA/literacy and mathematics, tasks should allow for more than one valid interpretation or viewpoint; for example, it is the quality of support that is marshaled in support of a position, not the particular position taken, that is important to the success of tasks asking for written argumentation in ELA or mathematical arguments. Multiple viable arguments should be possible based on the assignments and stimulus information provided in each performance task.
- **Require scoring that focuses on the essence of the task** (Dana and Tippins, 1993; Kane and Mitchell, 1996). Scoring of student work from a task will use multiple rubrics uniquely matched to the primary content claims and targets identified by the task developer in the task forms.
- **Be feasible for the school/classroom environment** (Dana and Tippins, 1993). Performance tasks are constructed so they can be delivered effectively in the school/classroom environment. Considerations for task specifications include, but are not limited to, student-teacher interactions, materials and technology needs, and allotted time for assessment. Performance tasks will adhere to a framework or specifications that can be used
    - by item writers to develop new tasks that focus on different content but are comparable both qualitatively and statistically, and
    - by classroom teachers for creating assessments providing meaningful evidence for both formative and summative purposes.

General Specifications for Smarter Balanced Performance Tasks

Specifications presented in this section pertain to all Smarter Balanced performance tasks. To avoid unnecessary repetition, however, cross-component specifications that have been developed by groups other than the Performance Task Work Group are not presented here, even though those specifications are relevant to performance tasks. Such cross-component specifications include specifications for stimulus materials for all items and performance tasks, specifications for scoring rubrics, and guidelines for formatting and style, bias and sensitivity, and accessibility and accommodations. However, if additional specifications are required in these areas for performance tasks, they are included. Similarly, unique specifications in these areas for ELA/literacy and mathematics or for different grade levels are addressed in the detailed target-specific ELA/literacy and mathematics specifications and only minimally here.

Overall structure of tasks

All Smarter Balanced performance tasks will consist of three basic components: stimulus presentation, information processing, and scorable product or performance. Information processing means student interactions with the stimulus materials and their content. It could include note taking, data generation, and any of a number of other activities that advance the students' understanding of the stimulus content or the assignment. All activities within a task must have a reason for being there (e.g., to increase understanding, for scaffolding, as early steps in product production, or for product creation itself). While a task will involve multiple parts and associated responses or products, the parts should be related, contributing to a whole or leading to the final

product or performance. More detail on the possibilities within the three basic task components is presented in specifications for ELA/literacy and mathematics performance tasks.

Task administration/setting

All tasks will be administered in controlled classroom settings with time limitations established by grade and subject. Smarter Balanced item and task specifications assume computer delivery of the items and tasks. Most tasks will be long enough to warrant two test administration sessions. Such sessions will be same-day, back-to-back sessions with short breaks in between sessions. Time requirements for tasks are provided in subject-specific performance task specifications.

Introduction of tasks to students (Stiggins and Chappuis, 2005; Khattri, Reeve and Kane, 1998)

Student directions for all tasks will begin with an overview of the entire task, briefly describing steps in all sessions. The overview should give the students advanced knowledge of the scorable products or performances to be created.

Allowable teacher-student interactions

Teachers roles during the administration of performance tasks may vary by task. Allowable teacher-student interactions for a task will be standardized (i.e., carefully scripted or described in task directions for purposes of both fairness and security). Teachers are not to offer assistance to students in the production of their scorable products or presentations.

Group work

In the interest of fairness and security, small group work by students during performance tasks will generally not be allowed. The exception will be an occasional mathematics task during which student teams may be asked to perform a function such as generating data or engaging in other activities that in no way can advantage members of one group over another when they break up to do their individual work leading to scorable responses or products.

Organization of complex task directions

Dense text will be avoided. To the extent possible, directions, requirements, and needed information or guidance will be provided via bullets, tables, etc., and will be presented separately by task step or part.

Vocabulary

Definitions of specialized terminology/vocabulary that students are not expected to know will be provided within the text of task directions.

Stimulus complexity

Text readability/complexity (including scripts of video stimuli) should be at least one grade below grade level if reading is not being assessed. This is not to say that students should not be reading at grade level; readability is being controlled this way so that the students can stay focused on the task

and the content and skills other than reading that are being measured. More information on allowable stimulus materials is provided in separate stimulus specifications.

Access to stimuli

Because performance tasks involve multiple texts and/or other resources that students need to access at various stages as they prepare for and create their final products/presentations, icons or "buttons" on the computer screens will make these resources available in an efficient, user-friendly format. For some tasks, there may still be a need for paper-based stimuli, note paper, or other materials. If hardcopy materials are collected in between sessions, at the start of a new session, each student must be given the same set he/she used during the previous session.

Internet access

Because of concerns regarding security, time, and Internet capacity, students will not be given free access to the Internet for searches. Screen shots of search results and simulated searches can be incorporated into tasks.

Tools and other resources

To the extent possible, required tools and other resources will be provided to students within the online testing system. On occasion, teachers may have to assure student access to some items that should be readily available to them in the classroom or online.

Student work to be evaluated

All scorable products or performances will reflect individual student work. Every task will lead to multiple scorable products or performances or to a product or performance with multiple scorable components or attributes (Dunbar, Koretz and Hoover, 1991). Included in the task directions will be information for the students about what work will be scored. Nothing more will be said about selected-response and constructed-response questions. For work for which general rubrics will be used in scoring, such as essays and oral presentations, the students will be informed about what attributes of their work will be scored (Stiggins and Chappuis, 2005).

Scoring requirements (Dunbar, Koretz and Hoover, 1991)

An extensive amount of student effort and work reduced to a limited number of points in scoring has no greater measurement value or reliability than a single short item worth as many points. Therefore, to achieve score ranges for performance tasks that reflect the amount of work (and evidence) produced, every task will lead to scores for individual students with a total task score range equivalent to that of several constructed-response questions combined. (Further specification of point ranges is provided in the ELA/literacy and mathematics general performance task specifications.) For any particular task, this will be achieved by totaling points across scorable products, task parts/components, or product attributes, and not by a single holistic score. While some performance task products may be scored by computers, it is assumed that much of the student work from performance tasks will require human scoring either centrally or through distributed scoring.

Rubrics (Kane and Mitchell, 1996)

Separate, unique rubrics for products, task components, or product attributes will be used. The point ranges in rubrics will vary depending on the scope and importance of what is being scored. Rubrics will be consistent with Smarter Balanced general rubrics.

Task templates/forms

Smarter Balanced performance task developers will create tasks by completing item specification forms for performance tasks, inserting the required information (e.g., meta-data, directions, rubrics, etc.) in the appropriate fields.

More information on the characteristics of and requirements for Smarter Balanced performance tasks is provided in subject-specific performance task specifications.