Race to the Top Assessment Program Spring 2013 Technical Review for the Partnership for Assessment of Readiness for College and Careers (PARCC)
July 3, 2013

Introduction
In April 2013, the U.S. Department of Education (Department) created and convened the Race to the Top Assessment Technical Review as part of the Department's continuing work to support the two consortia of states developing next generation assessment systems, the Partnership for Assessment of Readiness for College and Careers (PARCC) and the Smarter Balanced Assessment Consortium (Smarter Balanced). The review provided the consortia with an opportunity to share technical documentation and items being developed with external experts in assessment design and development, educational testing validity, accessibility, psychometrics, and English language arts and mathematics content.[1]

The Department appreciates that the consortia are in the middle of their assessment development. As a result, much of the data and documentation to confirm the technical quality of the assessment system is not yet available, and the consortia will continue to revise and improve their development processes over the coming months. The review focused on two broad areas of assessment development: (1) the consortium's research confirming that assessment results will be valid for their intended uses and (2) the test development process, including reviewing a sample of items and tasks. The experts' individual analyses provided information that both the consortium and the Department can use during this critical period. We expect the Technical Review to help each consortium to identify whether it has sufficient processes in place, has documented its work and development decisions to date, and has sufficient research underway or planned to confirm the technical quality of its assessment system.

The Technical Review is one component of the Department's Race to the Top Assessment Program Review, which is the way we are working with both consortia to support and provide oversight of their work. The Department intends to convene experts periodically for additional review and comments. This review focused on the technical quality of the consortia's assessment development. It did not examine other critical components of developing a large, multi-state assessment system, such as procurement, project management, organizational effectiveness, technology development, and professional capacity, outreach, and engagement. These topics are addressed through other components of the program review; more information about the program review, including an analysis of the consortium's progress during the first two years of the grant, can be found in the annual consortium reports, available at www2.ed.gov/programs/racetothetop-assessment/performance.html.

About this Document
To assist the consortia in considering the reviewers' comments, the Department identified high-level topics that were discussed during the Technical Review and asked each consortium to respond, identifying actions that will strengthen its work. The Department prepared this document for the convenience of the reader based on the individual reviewers' comments and the responses from the consortium. For ease in understanding this document, the Department included a preliminary analysis for each major topic area (assessment development; accessibility and accommodations; and research and planning) based on the reviewers' individual analyses and the consortium's responses:
   o Generally on track – the consortium's progress is on track and of high quality, though a few areas may need additional refinement or attention.
   o Some aspects on track, other areas need additional focus – the consortium has made progress, though some areas need attention or refinement during development.

[1] The Department requested that the technical experts, each nationally renowned in their fields, provide individual feedback to the Department and the consortium. The Department did not seek consensus advice or recommendations.

o Needs attention – the consortium has made some progress but a number of key areas need attention.
o Needs urgent action – the consortium needs urgent action to address concerns in key areas of development.

The Department will use the reviewers' analyses to guide its work with each consortium. The Department may identify where areas the consortia need additional assistance or focus as they incorporate Technical Review feedback into their ongoing assessment development. As part of the annual program review process, the Department will provide another opportunity for the Technical Review experts to analyze the continued progress of the consortia and to provide feedback to the Department and the consortia. The Department will also incorporate the Technical Reviewers' feedback into ongoing program review routines, expanding attention to key areas of the reviewers' feedback during monthly calls, annual site visits, and targeted support.

<u>Assessment Development</u>
Assessment Development
Highlights of Individual Reviewer Comments
- Many of the reviewers found that the items and text passages they saw were generally high quality and aligned to the content standards they were intended to measure.
- Several reviewers reported that the consortium is using evidence-centered design with fidelity.[2]
- Multiple reviewers found that test design, items, and scoring rubrics appear likely to provide clear, valid information for students, including students scoring well or poorly. The consortium will need to confirm through research that the test can discriminate well.
- Reviewers recommended that the consortium ensure that rapid acceleration in item development necessary to meet the consortium's item development goals does not lower item quality.

Highlights of PARCC Response
- The consortium reported that it has implemented item quality measures to ensure it continues to produce high-quality items through the tight timelines. In particular, PARCC conducts weekly meetings with the item developers to discuss items still under development and provide timely feedback. PARCC state and technical leaders also thoroughly review the items contractors are using as exemplars for training item writers.

English Language Arts/Literacy Content
Highlights of Individual Reviewer Comments
- Many of the reviewers applauded PARCC's use of authentic texts and thoughtful combinations of texts. They highlighted that the consortium has effectively selected texts worth reading.
- Numerous reviewers appreciated that the consortium's questions require students to use specific evidence from the text to support their answers.
- Reviewers generally found the constructed response items (assessment prompts requiring students to write essays) to be strong.
- Reviewers encouraged the consortium to ensure that items with multiple parts have plausible wrong answers across the parts and to consider partial-credit options.
- Some reviewers urged PARCC to ensure that individual instances of the assessment contain items and passages of varied complexity.

---

[2] Evidence-centered design is a process by which educators and content experts define the specific evidence they would expect to see a student demonstrate if s/he has mastered the content standards the assessment is measuring. Based on this depiction of what would demonstrate mastery, item developers craft items intended to give students an opportunity to show their skills and abilities.

Highlights of PARCC Response
- The consortium reported that it is working with item development contractors to ensure that items with multiple parts have plausible wrong answers across the parts.
- PARCC also stated that it is discussing scoring options that would allow for partial credit in collaboration with the research working group and advice from the technical advisory committee (TAC).[3]
- The consortium responded that it developed a cognitive complexity framework to ensure that students see items and passages across the full range of complexity on PARCC assessments, which the consortium will further investigate using field test data.

Mathematics Content
Highlights of Individual Reviewer Comments
- Several reviewers appreciated the clarity of the assessment design documentation.
- Some reviewers encouraged PARCC to simplify some of the language in instructions and test questions and to expand its use of plausible real life contexts.
- A few reviewers encouraged PARCC to expand its use of technology in the items that model or simulate data beyond the sample items provided for the review.

Highlights of PARCC Response
- PARCC reported that it revised its style guide for item writers to clarify expected use of language.
- The consortium stated that 25 percent of items on the PARCC assessments will be technology-enhanced items and the consortium is working with item developers to encourage development of innovative items. The consortium anticipates expanded use of technology in the first operational year in the 2014-2015 school year and beyond.

Preliminary Department Analysis: Generally on track.

Next Steps: Overall, the review demonstrated that PARCC has done substantial positive work in the area of assessment development. Several reviewers saw evidence that PARCC has established and used a strong process for developing assessment items. Many reviewers also highlighted PARCC's thoughtful selection of meaningful text passages for the English language arts/literacy assessment and careful mathematics assessment design. The next step is for the consortium to continue building high-quality assessment items while accelerating the pace of production.

---

[3] Educators and content experts in PARCC states serve on PARCC working groups dedicated to key activities, such as research, technology, English language arts, and mathematics. In addition, the PARCC technical advisory committee (TAC) is a group of research, psychometric, and accessibility experts who provide ongoing feedback to the consortium. See www.parcconline.org/technical-advisory-committee for more information on the TAC.

Accessibility and Accommodations
   Highlights of Individual Reviewer Comments
- Some reviewers appreciated the draft accessibility and accommodations handbook and the consortium's stated commitment to universal design for learning.[4]
- Several reviewers saw evidence of attention to accessibility in the process of defining the evidence statements that the consortium used in test design and development.
- Multiple reviewers encouraged PARCC to expand its accessibility work into all phases of item development, including expanding accessibility and accommodations training for item writers and reviewers with the aim of improving overall quality.
- Reviewers emphasized that PARCC should carefully research the way items work for the full range of English learners and students with disabilities (i.e., examining how English learners at different levels of English proficiency interact with items of varying complexity; examining how students with disabilities with different disabilities and disabilities of distinct levels of severity perform on items at all levels of cognitive complexity).

   Highlights of PARCC Response
- PARCC responded that it plans to incorporate additional training on accessibility and accommodations for item writers during the second phase of its item development work (beginning fall 2013).
- To ensure that the consortium implements the most appropriate guidance regarding accommodations, PARCC also plans to research item function for students with disabilities through studies based on field test results and make adjustments as needed to its accommodations manual. As the consortium committed in its approved application, PARCC will need to research the way items function for the full range of English learners and students with disabilities.

Preliminary Department Analysis: Needs attention.

Next Steps: The consortium should increase its emphasis on accessibility and accommodations throughout assessment development by more deeply engaging experts on students with disabilities and English learners. PARCC should continue its plans to improve and increase accessibility training for item developers. In addition, the consortium will need to ensure that its research plans are sufficient to evaluate the validity and reliability of the assessments for the intended purposes, including researching the accessibility of the assessments for English learners of varying levels of English proficiency and students with disabilities experiencing disabilities of a range of types and severity.

Research and Planning
   Highlights of Individual Reviewer Comments
- Several reviewers appreciated that PARCC had asked its TAC to prepare "white papers" (short technical issue briefs or literature reviews) and urged expanded use of the research articulated through these papers.
- Most reviewers saw evidence that the consortium is planning for critical future research projects.
- Many reviewers strongly suggested additional near-term research to ensure that items and tasks work as intended; that students understand what they are being asked to do; and that test design, items, and scoring rubrics provide strong discrimination for students who score at high and low

---

[4] Universal design for learning is an approach to instruction and assessment that considers all students in development of tasks and assessment items with the goal of making content flexibly available to students with a range of needs. It functions when educators design materials from the beginning in a widely accessible manner while maintaining high expectations for all students, including students with disabilities and English learners.

levels as well as students who achieve mid-range scores. The reviewers recommend that the consortium incorporate information about whether students involved in those studies have yet received instruction on the material the assessment measures as a variable in the studies.

- Reviewers suggested that the consortium expand test security planning and training.
- Some reviewers urged PARCC to begin drafting report templates to ensure they are clear and concise.

Highlights of PARCC Response
- PARCC reported that it completed small-scale studies during item development in spring 2013 regarding how students interact with the assessment tasks, how students with disabilities respond to the assessment items, and which of two writing rubric models is more valid and reliable.
- During the summer of 2013, PARCC is doing four additional studies: (a) small-scale item tryouts including one-to-one interviews with students in which students are asked to "think aloud" as they solve the more complex mathematics items; (b) a narrative writing study to investigate whether student responses can be used to support both reading and writing scores; (c) a study of whether and how partial credit may be available on some English language arts/literacy items; and (d) a cognitive lab study to investigate students' interaction with mathematics and English language arts/literacy items on tablets and laptops. Additional research will draw on data from the field test, which will be administered in the spring of 2014.
- PARCC also reported that it drafted a preliminary item security policy for the summer 2013 studies and will have a Test Administration Manual, including test security information, available in draft form in August 2013. PARCC states have been contributing to the conversation about the test window by modeling test administration possibilities.

Preliminary Department Analysis: Some aspects on track, other areas need additional focus.

Next Steps: While high-level plans for research are generally on track, additional attention is needed in this area as PARCC implements those plans in both the near and long term comprehensively and with fidelity. PARCC should be responsive to the reviewers' recommendations regarding the timing and types of research needed. The next steps are for PARCC to revise its comprehensive research plans, in consultation with its TAC, adjusting those plans as necessary to best inform ongoing item development.