

National Assessment Governing Board

Assessment Development Committee

May 15-16, 2014

AGENDA

| Thursday, May 15, 2014 | | |
|-------------------------------|--|---|
| Noon – 4:00 pm | Closed Session ACTION: Science Interactive Computer Tasks | Secure material sent under separate cover |
| Friday, May 16, 2014 | | |
| 10:00 – 10:05 am | Welcome, Introductions, and Agenda Overview <i>Shannon Garrison, Chair</i> | |
| 10:05 – 10:25 am | Technology and Engineering Literacy (TEL) Assessment Update <i>William Ward, NCES</i> | Attachment A |
| 10:25 – 11:00 am | Transitioning to NAEP Technology-Based Assessments in Reading and Mathematics <i>Eunice Greer, NCES</i> | Attachment B |
| 11:00 – 11:25 am | Follow-up on NAEP Read Aloud Study <i>Grady Wilburn, NCES</i> | Attachment C |
| 11:25 – 11:50 am | Options for NAEP Assessments in U.S. History, Civics, and Geography <i>Mary Crovo, Governing Board Staff</i> <i>Eunice Greer, NCES</i> | Attachment D |
| 11:50 am – 12:10 pm | NAEP and Next Generation Science Standards: A Comparison Study <i>Teresa Neidorf, AIR</i> | Attachment E |
| 12:10 – 12:30 pm | Contextual Variables: How New Guidelines are being Implemented Across Subject Areas <i>James Deaton, NCES</i> | Attachment F |
| Information Item | NAEP Item Review Schedule | Attachment G |

Technology and Engineering Literacy (TEL) Assessment Update

Because of the growing importance of technology and engineering in the educational landscape, and to support America's ability to contribute to and compete in a global economy, the National Assessment Governing Board initiated development of the first national assessment in Technology and Engineering Literacy (TEL). In 2014, the TEL assessment, designed to gauge how well students can apply their understanding of technological principles to real-life situations, was operationally administered to approximately 22,000 grade 8 students in over 800 public and private schools.

In this session, NCES will share the latest anecdotes from students and school staff from the 2014 TEL administration including a discussion of three positive themes that emerged from the TEL administration debriefing:

- Levels of student engagement;
- Student reactions to TEL supplemental block
- Reactions to TEL administration from school staff.

Additionally, an update will be provided on the status of NCES web-based outreach efforts on the TEL information page (<http://nces.ed.gov/nationsreportcard/tel/>), currently hosted on the Nation's Report Card website. The TEL information page was designed to increase overall awareness about TEL and inform and prepare the public for the assessment. A key feature of the TEL information page is a sample TEL task entitled "Wells". At the last meeting, the Assessment Development Committee (ADC) reviewed preliminary plans for posting additional supporting materials regarding the Wells task. In this session, the update on these activities will include a discussion contrasting the current web-based outreach efforts from future TEL task release strategies. There will also be a discussion of how best to balance the need for web-based outreach and future task releases for TEL reporting, given program resource limitations.

UPDATE on the Transition to Technology-Based Assessments (TBA) for Reading and Mathematics

Transitioning reading and math paper-and-pencil tests to technology-based assessments (TBA) is among the major challenges and opportunities facing the NAEP program. As we have already discussed, the intent is not to simply transfer the existing items to electronic delivery, but to introduce new types of technology-enabled items that can measure knowledge and skills that could not be tested—or could not be tested as well—on paper. Given the central role that trend plays in NAEP assessments and reporting, there is also a desire to continue to measure progress within each subject area by maintaining NAEP trend lines despite the change from paper-and-pencil to technology-based assessment modes.

At this session, NCES will present an update on our work to transition NAEP reading and mathematics assessments to technology-based assessments. Both the mathematics and reading teams are working to create the transition items for the 2015 TBA Start-Up/Bridge Study as well as the new items that will make up the 2016 and 2017 pilot assessments.

Since the March Governing Board meeting the focus of the 2015 mathematics TBA Start-up development continues to be on transitioning as many of the existing paper-and-pencil items as possible, including the existing constructed-response items, to new TBA item types. Work on the new development for the 2017 pilot has included conversations with the Mathematics Standing Committee on potential scenarios for a small number of grade 4 and grade 8 extended tasks being planned for 2017.

In reading, NCES convened an ad hoc committee of experts to review and discuss early results from usability studies. These studies are conducted to help us understand how students understand and respond to our technologies, and what we can do to improve performance. The ad hoc committee heard and discussed results of early studies in reading that were focused on passage layout and formats for item presentation. In the May meeting, NCES will also describe the membership of the Ad Hoc Committee on Reading TBA, present results from the usability studies and re-cap the ad hoc committee's recommendations.



Follow-up to NAEP Read Aloud Study

The administration of a read aloud provision on reading comprehension assessments is a topic of much discussion among educators and policy makers. Currently, states and districts follow various practices regarding the read aloud provision. Some jurisdictions designate read aloud as a modification (*does* violate the reading construct), while others label read aloud as an accommodation (*does not* violate the reading construct).

According to the Board-adopted NAEP Reading Framework, the assessment measures reading comprehension by requiring students to read passages of written English text and to answer questions about what they have read. Based on this defined construct for the assessment, reading aloud the NAEP reading passages would be a modification and is not allowed on the NAEP reading assessment.

To provide data to inform the read aloud discussion, NCES contracted with Dr. Jamal Abedi of the University of California at Davis, to conduct a study of the read aloud provision. The study was titled, “Effectiveness, Validity, Differential Impact, and Feasibility of Read-Aloud Accommodations in NAEP Reading Assessment”. In this study, Dr. Abedi explored the use of the read aloud accommodations on NAEP. This was investigated by administering the NAEP reading assessment to 4th and 8th grade students in California and placing them in one of three conditions: read aloud the passages and items; read aloud everything except the reading passages; and no read aloud.

At the February 28, 2014 Assessment Development Committee meeting, NCES presented the findings of Jamal Abedi’s read aloud study. These findings suggested that while the read aloud everything (passages and items) might have negative measurement implications, that reading aloud everything *except* the passages might be an acceptable accommodation, particularly in 4th grade.

The presentation NCES will give to the Assessment Development Committee will focus on ways that the study findings might be interpreted, and plans for next steps.

National Assessment Governing Board Assessment Development Committee

Report of February 27-28, 2014

Excerpt

Results of the NAEP Read Aloud Study

Grady Wilburn of NCES provided a briefing on a study of various read aloud conditions. The principal investigator for the study was Jamal Abedi, of the University of California at Davis. NCES commissioned the study to examine the utility of read aloud on NAEP. Research questions included:

1. Does the read aloud help students with disabilities (SD) and English language learners (ELL)?
2. Does the read aloud help the non-target group (non-SD/ELL students)?
3. Does the accommodation help the target group more than it helps the non-target group?

The study also was intended to contribute to the research on the read-aloud provision.

Mr. Wilburn described the sample of students in 4th and 8th grade and the materials they were provided in the study. Students were randomly assigned to one of several experimental conditions: read aloud all (passage, directions, items); read aloud everything except the passage; and no read aloud. All groups of students (SD, ELL, and Non SD/ELL) received each treatment.

ADC members were presented with findings of the study in terms of 4th and 8th grade student performance in the various experimental conditions. Results showed that the read aloud was effective at grade 4, but had mixed results at grade 8. The read aloud was not shown to be valid across all conditions in grades 4 and 8.

The ADC engaged in a lengthy discussion of the study methodology, findings, and conclusions. First, the study labeled read aloud as an accommodation. However read aloud is a modification of the Board-adopted Reading Framework definition of reading comprehension. According to the Framework, the NAEP Reading Assessment requires students to read passages of text written in English and respond to questions about what they have read. The NAEP Assessment is a reading comprehension assessment; it is not an assessment of listening comprehension.

ADC members noted that variability due to the human reader was not accounted for in the study. This variable could have significantly influenced the results. There was also a question about the sample size and generalizability of the findings. Particularly in the Specific Learning Disability subgroup, there are many types of learning disabilities. It was not clear how these were represented in the subgroup of students in the study. Another question that was not addressed in the study was how the test administrators were instructed to respond to a student who wanted them to re-read a part of the text to support an answer. Many other concerns about this study were expressed by the ADC.

In terms of implications, members noted that reading aloud a reading comprehension test to students sends a message to teachers of struggling readers at the early elementary level. Many of those students can learn to read, but individual students may learn to master various reading skills at a different pace. It would be inappropriate for NAEP to signal that struggling readers not receive reading instruction in the elementary grades. ADC members did acknowledge that a very small percentage of students do have disabilities that prevent them from reading written text.

ADC members thanked Mr. Wilburn for briefing the Committee on the study. The ADC found the study design and results interesting. However, the Committee reaffirmed that reading aloud reading passages on the NAEP Reading Assessment was not consistent with the construct being measured. The read aloud study did not support a change in the NAEP Reading Framework.

Transition to Technology-Based Assessments (TBA) for Civics, U.S. History and Geography

Transitioning NAEP's civics, U.S. history, and geography paper-and-pencil tests to technology-based assessments (TBA) is the newest challenge NCES will take up in the program-wide transition to TBA. As with the other subjects, the intent is not to simply transfer the existing items to electronic delivery, but to introduce new types of technology-enabled items that can measure knowledge and skills that could not be tested—or could not be tested as well—on paper.

At this session, NCES staff will begin with a review of the legislation that called for the development of the assessments as well as the history of the frameworks and assessments for civics, U.S. history and geography. The TBA environment removes form and booklet constraints that were part of the paper-and-pencil culture. These new freedoms give us the opportunity to re-conceptualize a more efficient assessment design including how we group and use items.

NCES will present our proposed design for transitioning NAEP to technology-based assessments according to the current schedule. The TBA start-up is planned for 2016; pilot of new items in 2017; and the first operational assessment is scheduled for 2018. Additionally, as with other subjects we have already discussed, ideas for newer item types will be introduced.



NCES Plans for Comparing NAEP and the Next Generation Science Standards

Prepared by AIR for NCES – April 18, 2014

Over the past few years, new national standards documents have been developed in the areas of science, technology, engineering, and mathematics (STEM). The first of these STEM standards documents were the *Common Core State Standards in Mathematics* (CCSS-M) developed by a state consortium and published in 2010. In 2012, the National Research Council (NRC) published *A Framework for K-12 Science Education* that provided a vision for K-12 science education. This guiding document was the basis for the state-led development of the *Next Generation Science Standards* (NGSS) that were published in 2013. Most recently, the NRC also published *Developing Assessments for the Next Generation Science Standards* in 2014. These national documents are leading to major changes in state curricula and assessments.

To inform ongoing discussions of NAEP's role in emerging national systems of large-scale assessments, NCES is planning a comparison study of NGSS with the NAEP assessments in science, technology and engineering literacy (TEL) as well as relevant aspects of the mathematics assessment. The purpose of the study is to provide evidence of where NAEP can provide good assessment models and national measures of the student understandings described in the K-12 Framework and the NGSS. In addition, the study will identify areas where NAEP provides additional measures that augment what can be assessed by states based on the NGSS. The results of the study will be important as NAEP continues to explore new approaches to STEM assessment designs and recommendations from the Future of NAEP summit. There are two levels to be considered when comparing NAEP with NGSS/NAEP – i) a framework comparison and ii) an item pool comparison.

A framework comparison study is planned for 2014. The intent of this study is to determine the extent to which the STEM frameworks in NAEP are intended to measure the content and scientific and engineering practices in the NGSS. Using methodology successfully applied in previous framework comparison studies between NAEP and international assessments in science and mathematics, this study will compare the description of content and practices to be assessed in the NAEP Science and TEL Frameworks at grades 4, 8 and 12 with the performance expectations in the NGSS at the corresponding grade levels. In addition, portions of the NAEP Mathematics Framework (particularly content areas of measurement and data analysis, statistics, and probability) will be compared with expectations for the practice of using mathematics and computational thinking as well as connections to the CCSS-M specified in the NGSS. The comparison will also consider the alignment between NAEP frameworks and the larger set of guidelines related to science and engineering practices, disciplinary core ideas and crosscutting concepts in the original K-12 framework document that was the basis for the development of the NGSS. The results of the study will provide an evaluation of the extent of overlap of the NGSS and NAEP assessment frameworks as well as a description of specific areas where the NAEP frameworks and NGSS cover similar or different content and skills.

If resources allow, a second phase of the study will be conducted in 2015 to compare items and tasks in the NAEP STEM assessments with NGSS. The timing of this study takes several factors into consideration. First, by 2015 the technology-based NAEP science item pool (discrete items, interactive computer tasks and hybrid hands-on performance tasks) as well as TBA items in NAEP mathematics will be developed. Second, results from the 2014 TEL assessment will be available to inform the evaluation of TEL items and tasks. Third, the nature and scope of the item pool comparison study will be informed by the results of the framework comparison conducted in 2014. In addition, further advancements in the application of NGSS to the development of state science curricula and assessments will be considered in the design and reporting of results from the study.



Contextual Variables: How New Guidelines are being Implemented Across Subject Areas

The “Policy Statement on NAEP Background Questions¹ and the Use of Contextual Data in NAEP Reporting” was unanimously adopted by the National Assessment Governing Board in August 2012. This statement recommended eleven implementation guidelines covering three main areas: 1) *Questions and Questionnaires*, 2) *Data Collection*, and 3) *Reporting*. The guidelines were adopted following recommendations by an expert panel (chaired by Marshall Smith), a wide range of public comments, and feedback from NCES.

At the May 2014 meeting, NCES will discuss how many of these new guidelines have been implemented and plans for future development cycles within the subject-specific contextual questionnaires. This presentation will be focused on implementing applicable guidelines across subject areas. Therefore, the presentation will concentrate on implementation guidelines for *Questions and Questionnaires* and *Data Collection*.

In this session, NCES will:

- discuss shifting the questionnaire design approach (i.e., one that has historically favored breadth of coverage over depth of coverage) to one that is more balanced and utilizes clusters of questions
- detail efforts to eliminate duplicative or low-priority questions
- spotlight where NCES plans to include contextual questions from international assessments
- examine considerations for preserving trend in relation to transitioning subjects to technology-based assessments; and
- report how considerations for overall timing of questionnaires and potential for spiraling questions will be positively impacted by switching to technology-based assessments and by creating clusters of questions around the same topic.

¹ Now referred to by the Board as “Contextual Questions.”



**Assessment Development Committee
Item Review Schedule
December 2013 – August 2014**
(Updated 4/18/14)

| Review Package to Board | Board Comments to NCES | Survey/Cognitive | Review Task | Approx. Number of Items | Status |
|-------------------------|------------------------|------------------|---|-------------------------|----------------------|
| 11/21/13 | 12/13/13 | Survey | 2015 Operational Reading (4, 8) | 78 | ✓ |
| 11/21/13 | 12/13/13 | Survey | 2015 Operational Math (4, 8) | 90 | ✓ |
| 1/31/14 | 2/26/14 | Survey | Reading question pool* (4, 8) | 78 | ✓ |
| 2/13/14 | 3/6/14 | Cognitive | 2015 Operational Reading (4, 8) | 55 | ✓ |
| 2/13/14 | 3/6/14 | Cognitive | 2015 Operational Math (4, 8) | 91 | ✓ |
| 3/21/14 | 4/2/14 | Cognitive | 2015 Operational Reading (12) | 17 | ✓ |
| 4/11/14 | 4/25/14 | Survey | 2015 Operational Science (4/8/12) | 139 | ✓ |
| 5/1/14 | 5/22/14 | Cognitive | Science ICT Beta Reviews (4, 8, 12) | 6 | For review on May 15 |
| 7/17/14 | 8/7/14 | Cognitive | Science ICT Clearance Review (4, 8, 12) | 15 | |
| 7/17/14 | 8/6/14 | Cognitive | 2015 Operational Science (4, 8, 12) | 96 | |

NOTE: Alpha builds will be presented to the ADC during their in-person and virtual meetings. These will not be submitted before the review. NCES will submit the beta builds to the ADC prior to their review meetings. (Alpha and beta builds are the first- and second-draft versions of the rendered task, respectively.)

*A survey question pool represents all the questions that have been administered in operational assessments.