National Assessment Governing Board

Assessment Development Committee

November 20, 2015

AGENDA

10:00 – 11:00 am	**Closed Session** **Joint Meeting with Reporting and Dissemination (R&D) Committee** Plans for Reporting the 2014 Technology and Engineering Literacy (TEL) Assessment Results 　　*Rebecca Gagnon, R&D Chair* 　　*Shannon Garrison, ADC Chair* 　　*Daniel McGrath, NCES* 　　*Robert Finnegan, ETS*	Attachment A
11:00 – 11:05 am	**Assessment Development Committee** Welcome, Introductions, and Agenda Overview 　　*Shannon Garrison, Chair*	
11:05 – 11:30 am	Update on NAEP Digital-Based Assessments 　　*Eunice Greer, NCES*	Attachment B
11:30 am – 12:30 pm	**Closed Session** Additional Analysis of NAEP 2015 Reading and Mathematics Data: Focus on Frameworks and Items 　　*Andreas Oranje, ETS*	Attachment C
	Information Item: NAEP Item Review Schedule	Attachment D

NAEP Technology and Engineering Literacy (TEL) Assessment Update

Joint Meeting of the Reporting and Dissemination R&D Committee and Assessment Development Committee (ADC)

Purpose

The purpose of the November 20, 2015 joint R&D and ADC meeting is to provide an update on the TEL Report Card website development and an overview of the website review process. The two Committees will be briefed on changes made to the TEL website design since the March 2015 joint R&D and ADC meeting. In addition, Board members will discuss overarching ideas on the website's appearance and accessibility, types of findings to highlight, and other issues for a successful release of the TEL results.

Background on TEL

In 2005 the National Academy of Engineering and the National Research Council called on the Governing Board to add a new NAEP assessment in the area of Technological Literacy. The Board extensively deliberated this recommendation and gathered broad-based feedback before deciding to add this assessment to the NAEP schedule. Subsequently, the Board conducted a multi-year, comprehensive framework development process involving thousands of educators, policymakers, IT professionals, engineers, business representatives, testing experts, and others.

Eventually renamed Technology and Engineering Literacy, or TEL, this innovative assessment was based on a Board-adopted framework that called for a unique combination of scenario-based tasks and discrete test questions, all of which were to be administered via a computer-based platform. After various stages of test development and full-scale pilot test, the TEL assessment was administered in spring 2014 to a nationally representative sample of more than 20,000 8th graders. The assessment was designed to measure how well students can apply their understanding of technology principles to real-life situations. Results will be available at the national level only and will be released as The Nation's Report Card.

TEL Report Card Site to Date

At the March 2015 Board meeting, the R&D and ADC met jointly to receive a preview from National Center for Education Statistics (NCES) on the TEL Report Card design plans. The Committees offered a range of recommendations for the site design under the heading of three major goals:

- Enable users to explore the scenario-based TEL tasks;

- Provide results of interest to policymakers, educators, and the general public; and

- Use engaging multi-media features on the homepage to explain more about what the TEL assessment measures.

The ADC convened in August 2015 to view screen shots of the updated Report Card site design. The ADC focused on ensuring the prominence of the TEL tasks and content on the site and making that information as user-friendly and accessible as possible. The ADC was pleased that many R&D and ADC recommendations from the March 2015 joint meeting were implemented in the updated design of the Report Card site.

Update on the TEL Report Development and Release

The TEL Report Card is scheduled for a spring 2016 release, pending finalization of achievement levels by the Board in November 2015. NCES is incorporating improvements requested by the Board and integrating the TEL Report Card into the larger NAEP Report Card site. The Report Card home page will include main results, as well as motion graphics and prominent access to example task pages designed to provide an inviting and informative introduction to TEL for a broad audience. The example task pages will include narrated video walkthroughs of four scenario-based tasks, including exploration of the tasks, data on student performance on the tasks, and data on related survey questionnaire items. The first fully-functional version of the Report Card website is scheduled for NCES review in December.

Update on NAEP Digitally-Based Assessment Development and Observable Data

NAEP's transition from a paper and pencil assessment began in 2011 with the laptop-based assessment of Writing at grades 8 and 12, and has expanded annually to include TEL, grade 4 Writing, and most recently Science, Mathematics and Reading. NAEP's decision to transition to digitally-based assessment (DBA) is driven by its goals to align more closely to emerging trends in educational practices in schools; to lead and inform large-scale assessment's transition to computer-based assessment models; and to take advantage of the affordances of technology as they relate to opportunities to collect and use "observable data" to better understand student's performances on NAEP.

The 2015 administration of NAEP's Reading and Mathematics items in the digitally-based assessment mode included the collection of a number of pieces of "observable data". Work has begun that explores whether patterns in the observable data are associated with performance on NAEP, and what we can learn from the observable data about what students actually do as they move through the assessment. Results from these important lines of inquiry will inform reporting and interpretation of results from NAEP's digitally-based assessments.

This session will address four questions associated with NAEP's collection of observable data:

1. What is meant by "observable data"?
2. What types of observable data are currently being collected as part of NAEP digitally-based assessments?
3. What are some examples of observable data, and how can we use them to more fully understand students' performance on NAEP?
4. Are there limitations to the collection of observable data, and what does this mean for NAEP?

Digitally-based Assessments and NAEP Trend

Overview

As NAEP transitions from paper to digitally-based assessments (DBA), an important question is how this transition affects trend reporting. To address this question NCES and NAEP contractors have done the following:

(1) Designed, implemented, and extended bridge studies to investigate the effect of mode changes on score distributions;

(2) Developed a decision tree to describe the key factors for subsequent analysis and decision making about trend reporting.

The closed session presentation to the Assessment Development Committee (ADC) will focus on the extended bridge studies component. Specifically, the ADC will receive a briefing on initial results comparing paper and tablet based versions of the Reading and Math assessment items administered as part of the 2015 NAEP. A briefing to the Committee on Standards, Design and Methodology (COSDAM) will focus on technical aspects of the two activities outlined above.

Extended Bridge Studies

Two bridge studies have been planned, one of which is currently being executed. Data collection for the first bridge study was part of the 2015 operational administration and entailed national samples in all three grades for Math, Reading, and Science. In these samples, a tablet-based version of the various NAEP instruments was administered on NAEP-provided tablets and analysis is currently under way. The goal is to compare the results from these digitally-based assessments to the paper-based assessments. The second bridge study currently planned would occur in 2017 in Math and Reading in 4th and 8th grade and entails small state-level samples participating in the paper-based assessment alongside larger state-level samples participating in the tablet-based assessment. Some tentative sample sizes have been proposed for the various components and are under discussion. The goal of this second study would be to (a) look at the stability of the mode differences (if any) across years (2015 and 2017) and (b) to estimate mode differences at the state level.

Alignment of NAEP Mathematics Items to the Common Core State Standards (CCSS)

The NAEP Validity Studies (NVS) panel conducted a study of the alignment between the 2015 NAEP item pools for grades 4 and 8 and the CCSS in mathematics. The study was led by Daro and Hughes, who also directed the earlier NVS study comparing the NAEP Mathematics Framework and the CCSS in mathematics. The results are intended to provide context for the interpretation of the 2015 NAEP results.

The central research questions for the study are:

- To what extent does the NAEP item pool include content that is not targeted by the Common Core State Standards for instruction at or before the grade level tested by NAEP?
    o What is the alignment profile for NAEP to CCSS across NAEP domains?

- To what extent do the Common Core State Standards target for instruction, at or before the grade level tested by NAEP, content that is not assessed by NAEP?
    o What is the alignment profile for CCSS to NAEP across CCSS domains?

On September 16-17, AIR convened a panel of 18 mathematicians, teachers, supervisors of curriculum, and mathematics educators to review all items in the 2015 NAEP mathematics item pools for grades 4 and 8. After dividing into grade groups (nine individuals for each grade), panelists classified each NAEP item into a CCSS standard or indicated that the item did not fit with any CCSS standard. Panelists worked independently to classify each item but then met as a group (within grade level) to discuss items for which the panelists were split regarding the central distinction needed to answer the study questions: whether or not the item should be classified into a standard at or below the grade level assessed by NAEP. After discussion, panelists were free to change their initial classification or not, as they desired.

When multiple standards could be applied to the same item, the directions to panelists helped them to standardize their choices. For example, panelists were instructed to start searching for matches among the CCSS standards at the grade level of the NAEP assessment. If no matches were found, they were instructed to search backward to earlier grades, and if that still did not produce a match, they were to search forward to later grades. Another guideline instructed them to treat with equal weight the possibility that the item did fit into a standard or did not fit into any standard. That is, they were not to default to one or the other of these choices when in doubt, but to make their best judgment.

A subset of NVS panel members reviewed and critiqued the report prior to its release on October 26, 2015. The report is available through the NVS page on the AIR website at
http://www.air.org/project/naep-validity-studies-nvs-panel

Assessment Development Committee
Item Review Schedule
November 2015 – August 2016
10/19/2015

Review Package to Board	Board Comments to NCES	Survey/ Cognitive	Review Task	Approx. Number Items	Status
11/20/15	12/16/15	Cognitive	2019 Math (4, 8) Pilot (SBT) Draft builds	4 tasks	
1/8/16	1/22/16	Cognitive	2019 Reading (8) Pilot (SBT) Draft builds	2 tasks	
1/25/16	2/16/16	Cognitive	2019 Reading (12) Pilot (SBT) Concept sketches	2 sketches	
2/23/16	3/11/16	Cognitive	2019 Math (4, 8) Pilot (SBT)	4 tasks	
2/24/16	3/11/16	Cognitive	2019 Math (12) Pilot (SBT) Concept sketches	3 - 4 sketches	
4/14/16	5/6/16	Survey	2018 Social Studies (8) Pilot (US History, Civics, Geography)	130-140	
4/14/16	5/20/16	Survey	2019 Math (4, 8) Pilot	20	
4/14/16	5/20/16	Survey	2019 Reading (4, 8) Pilot	10	
4/28/16	5/20/16	Cognitive	2018 Civics (8) Pilot (DI)	80	
4/28/16	5/20/16	Cognitive	2018 Geography (8) Pilot (DI)	80	
4/28/16	5/20/16	Cognitive	2018 US History (8) Pilot (DI)	80	
4/28/16	5/20/16	Cognitive	2019 Reading (4) Pilot (SBT) Draft builds	2 tasks	

Review Package to Board	Board Comments to NCES	Survey/ Cognitive	Review Task	Approx. Number Items	Status
4/28/16	5/20/16	Cognitive	2019 Reading (8) Pilot (SBT)	2 tasks	
4/28/16	5/27/16	Survey	2017 Math (4, 8) Operational	130-150	
4/28/16	5/27/16	Survey	2017 Reading (4, 8) Operational	130-150	
4/28/16	5/27/16	Survey	2017 Writing (8) Operational	130-150	
6/2/16	6/16/16	Cognitive	2019 Reading (12) Pilot (DI) Passage Review	4 passages	
7/20/16	8/12/16	Cognitive	2017 Reading (4, 8) Operational (DI)	25-30	
7/20/16	8/26/16	Cognitive	2019 Reading (4) Pilot (SBT)	2 tasks	
7/20/16	8/12/16	Survey	2019 Science (4, 8, 12) Pilot Initial item pool review Tentative	100-110	
7/21/16	8/12/16	Cognitive	2017 Math (4, 8) Operational (DI)	210	

"SBT" indicates Scenario-Based Tasks
"DI" indicates Discrete Items