

National Assessment Governing Board

Assessment Development Committee

Report of February 29 and March 2, 2012

February 29, 2012

Closed Session 9:00 a.m. – 3:00 p.m.

In accordance with the provisions of exemption (9)(B) of Section 552b(c) of Title 5 U.S.C., the Assessment Development Committee (ADC) met in closed session on February 29, 2012 9:00 a.m. to 3:00 p.m.

Attendees: ADC – Alan Friedman (Chair), Susan Pimentel (Vice Chair), Shannon Garrison, Doris Hicks, Brent Houston, Hector Ibarra, Dale Nowlin, Cary Sneider; Governing Board Staff – Mary Crovo, Michelle, Blair; AIR – Kim Gattis; ETS – Greg Vafis, Lonnie Smith.

NAEP Civics Cognitive Pilot Items

The Assessment Development Committee (ADC) began its closed session with a review of secure Civics pilot items at grades 4, 8, and 12. These items are proposed for pilot testing in 2013, in preparation for the 2014 operational assessment. Members commented that the questions were engaging and challenging at all grade levels. A number of questions contained original documents, photographs, political cartoons, or data presented in chart or graph form. Many items that used original source material required students to analyze and interpret the content and provide evidence to support their reasoning. In addition, members commented that many multiple choice questions tested sophisticated concepts beyond factual recall. While ADC had many comments on improvements to the items, overall members were very pleased with the breadth and quality of the Civics pilot items at each grade level.

NAEP Technology and Engineering Literacy (TEL) Tasks

In the second portion of the closed session, ADC members reviewed secure computer-based tasks being developed for the 2014 grade 8 Technology and Engineering Literacy (TEL) assessment. Lonnie Smith of ETS presented the current development timeline and an update on work that has occurred since the ADC last met in December 2011.

Mr. Smith noted that 21 interactive computer tasks are being developed for the 2013 TEL pilot test representing all three content areas of the TEL Framework: Design and Systems; Information and Communication Technology; and Technology and Society. Draft versions all tasks have been completed, although work is underway to gather input from external reviews. There will be a small-scale tryout of the tasks in March and April of 2012, with further task revisions based on the tryout. In May the ADC will engage in

a hands-on independent review of the completed tasks and items. Final ADC approval of all TEL tasks and items will occur at the August 2012 Board meeting.

Mr. Smith then presented the ADC with updated TEL tasks for review and discussion. ADC members were informed of ways in which their comments from the December 2011 review had been incorporated into each of the revised tasks. Members expressed their enthusiasm for the TEL tasks and commented that the graphics, sequencing, and content of the tasks were excellent. Members had a number of comments on the tasks, but most comments were in the category of “fine tuning” the directions or graphics for clarity.

February 29, 2012

Open Session

3:00 – 4:45 p.m.

Attendees: ADC – Alan Friedman (Chair), Susan Pimentel (Vice Chair), Shannon Garrison, Doris Hicks, Brent Houston, Hector Ibarra, Dale Nowlin, Cary Sneider; Governing Board Staff – Mary Crovo, Michelle, Blair; AIR – Kim Gattis; ETS – Greg Vafis, Lonnie Smith, Donnell Butler; Optimal Solutions Group – Linda Hamilton.

NAEP U.S. History, Civics, and Geography Pilot Background Questions

During the open session, the ADC reviewed pilot background questions in U.S. History, Civics, and Geography for the 2013 pilot test, in preparation for the 2014 assessments. While there were a number of questions in the review materials, there was substantial overlap among the social studies subject area questions. One key issue addressed in the ADC discussion related to how the term “computer” is defined and interpreted in the student background questions, particularly when there is an increasing array of digital devices that can be called a “computer.” ETS staff responded that alternate wording had been studied in cognitive labs and that new terminology would be used that was grade appropriate.

ADC members also discussed the desire to change and update certain questions to ask about more relevant information, and how these changes may affect trendlines. In some cases the ADC decided that it was more important to maintain trend, while in other cases they felt that revised questions were more appropriate to measure a given variable. A third issue discussed by the ADC related to the response scale associated with various questions. Members preferred the newer numerical version of the response scale (e.g., percentages of time spent on an activity) as opposed to the older version of questions that used a Likert-type scale to measure frequency (e.g., most of the time, all of the time, etc.). Changes were made to the background questions to reflect the ADC discussion in these and other areas.

The ADC also commented on improvements to the cognitive and background review booklets for online review in the future. It would be preferable to have embedded links in each question to view the pilot data (where available), rather than having separate pages in the PDF document. NCES and NAEP contractors will consider this format recommendation in planning for future item review documents.

March 2, 2012

Open Session

9:45 a.m. – 11:30 a.m.

Attendees: ADC – Alan Friedman (Chair), Susan Pimentel (Vice Chair), Shannon Garrison, Doris Hicks, Brent Houston, Hector Ibarra, Dale Nowlin, Cary Sneider; Governing Board Staff – Mary Crovo, Michelle, Blair; NCES – Suzanne Triplett; AIR – Kim Gattis, Fran Stancavage; ETS – Greg Vafis, Gloria Dion, Donnell Butler, Andreas Oranje; Pearson – Brad Thayer; Data Recognition Corp. – Paula Prah; HumRRO – Laurie Wise; Westat – Chris Averett.

NAEP Reading and Mathematics Pilot Background Questions

ADC members met in open session to discuss background questions for students, teachers, and schools in reading and mathematics. These items are proposed for pilot testing in 2014, in preparation for 2015 operational assessments at grades 4 and 8.

During the item review, ADC members noted that there are many more subject-specific questions for reading and math on the student background questionnaires than for the social studies subjects, particularly at 4th and 8th grades. This is partially due to the fact that in those grades students are mostly asked about social studies, as opposed to U.S. history, civics, or geography. Members noted that the more in-depth subject area questions provide a richer source of information on background variables.

As in the February 29, 2012 discussion on social studies background questions, ADC members commented on the use of the term “computer” and how that terminology should change going forward to be more inclusive of other types of digital tools. The issue of numerical response scales was also raised by the ADC. For some questions, ADC members offered alternate examples that they thought would be clearer to students when asking about instructional practices.

ACTION:

Under delegated authority from the Governing Board, the Assessment Development Committee approves the following cognitive and background questions, with changes to be communicated in writing to the National Center for Education Statistics:

- **Pilot Civics cognitive items in grades 4, 8, and 12 for the 2014 assessment**
- **Technology and Engineering Literacy (TEL) computer-based tasks in grade 8 for the 2014 assessment**
- **Pilot U.S. History, Civics, and Geography background questions for students, teachers, and schools in grades 4, 8, and 12 for the 2014 assessment**
- **Pilot Reading and Mathematics background questions for students, teachers, and schools in grades 4 and 8 for the 2015 assessment**

Briefing on Hewlett Foundation Automated Student Essay Scoring Prize

Mark Shermis, Professor at the University of Akron, provided a detailed briefing (via phone) to the ADC on this “X-prize” to develop software for scoring student responses to open-ended test question. Mr. Shermis the subject matter expert on the Hewlett competition and has considerable experience in automated essay scoring research.

The purpose of the competition is to determine if machine scoring can play a role in scoring new assessments being developed by the Common Core State Standards assessment consortia. Both PARCC and Smarter Balanced assessment consortia are collaborating on the Hewlett Foundation competition. The competition is currently in the first of three phases. In Phase 1, vendors were invited to use software to evaluate long student essays from six states. A total of 22,000 student essays were used in this phase of the competition. Phase 2 will explore how software can be used to evaluate students’ short answer responses and Phase 3 will address the scoring of responses to math questions.

Mr. Shermis provided detailed information on how this rigorous competition is organized and reported that nine vendors are competing in Phase 1. While the results of this phase are not yet public, Mr. Shermis reported on agreement statistics and other measures of quality of vendors’ scoring engines. He noted that many of the vendors have met the criteria established to measure agreement with the assigned scores on these various types of student essays. Complete findings from Phase 1 of the competition are scheduled for release in mid-April 2012.

ADC members engaged in a question and answer session to ask Mr. Shermis about additional details and findings of the competition. Members asked about the viability of automated scoring for the variety of NAEP questions. The ADC requested an update on the Hewlett competition at their May 2012 meeting.

March 2, 2012 Closed Session 11:30 a.m. – 12:15 p.m.

In accordance with the provisions of exemption (9)(B) of Section 552b(c) of Title 5 U.S.C., the Assessment Development Committee (ADC) met in closed session on March 2, 2012 11:30 a.m. to 12:15 p.m.

Attendees: ADC – Alan Friedman (Chair), Susan Pimentel (Vice Chair), Shannon Garrison, Doris Hicks, Brent Houston, Hector Ibarra, Dale Nowlin, Cary Sneider; Governing Board Staff – Mary Crovo, Michelle, Blair; NCES – Suzanne Triplett; AIR – Kim Gattis, Fran Stancavage, ETS – Greg Vafis, Donnell Butler, Andreas Oranje; ETS – Greg Vafis, Gloria Dion, Donnell Butler, Andreas Oranje; Pearson – Brad Thayer; HumRRO – Laurie Wise; Westat – Chris Averett.

Update on NAEP Mathematics Computer-Based Study

Gloria Dion of ETS provided a briefing to the ADC on a NAEP special study to examine the feasibility of computer-based adaptive testing. The study was conducted at the 8th grade as part of the 2011 NAEP assessment. Ms. Dion reported that the study design involved students being administered a set of multiple-choice math questions via computer. This “routing block” was broadly representative in terms of NAEP math content areas and contained a range of questions of varying difficulty levels.

The routing block items were scored “on the fly” and students were then presented with a second-stage block (hard, medium, or easy in item difficulty) depending on the student’s performance on the routing block. A control group of students were randomly assigned a second stage block. The student was designed to determine whether a two-stage computer adaptive approach would improve measurement, increase student engagement, and allow for meaningful interpretation of student performance across a wider range of student achievement.

Ms. Dion shared some preliminary findings from this special study during her presentation to the ADC. It appears from these early findings that there was an increase in measurement precision for the group of students who received a second-stage block based on performance in the routing block. Additional analyses are currently underway to examine factors related to student engagement and other data collected as part of this study. ADC members commented on the applicability to future NAEP assessments in mathematics and other subjects. The Committee requested an update on additional findings from the mathematics computer-based study at their May 2012 meeting.

I certify the accuracy of these minutes.



Alan Friedman, Chair

March 15, 2012

Date