

# **National Assessment Governing Board Committee on Standards, Design and Methodology**

**November 21, 2014**

## **Attendees**

**COSDAM Members:** Chair Lou Fabrizio, Vice Chair Fielding Rolston, Lucille Davy, James Geringer, Andrew Ho, Terry Holliday, and James Popham.

**Other Board Members:** Board Chair Terry Mazany.

**Governing Board Staff:** Executive Director Cornelia Orr, Michelle Blair, and Sharyn Rosenberg.

**Other Attendees:** Sue Betka, Acting Director of the Institute of Education Sciences and ex officio member of the Governing Board. NCES: Jing Chen, Patricia Etienne, Eunice Greer, Drew Malizio, Daniel McGrath, Bill Tirre, and Amy Yamashiro. AIR: Victor Bandeira de Mello, Markus Broer, Young Yee Kim, and Fran Stancavage. CRP: Edward Wofford. Education Week: Catherine Gewertz. ETS: Jonas Bertling, Steve Lazer, Rochelle Michel, Rebecca Moran, and Andreas Oranje. Hager Sharp: Lauren Werner. HumRRO: Laurie Wise. Optimal Solutions Group: Rukayat Akinbiyi. Pearson: Peg Heck and Paul Nichols. Westat: Keith Rust and Dianne Walsh. Widmeyer: Siobhan Mueller. Other: Andrew Kolstad.

## **Introductions and Review of Agenda**

Lou Fabrizio, Chair of the Committee on Standards, Design and Methodology (COSDAM), called the meeting to order at 10:30 a.m. and welcomed members and guests. Mr. Fabrizio noted that Mitchell Chester is a new member of COSDAM but unfortunately was unable to be present for this COSDAM meeting. Mr. Chester had a prior commitment out of the country that was made before he was notified of his appointment to the Board.

## **Project Update and Design Document for Technology and Engineering Literacy (TEL) Achievement Levels Setting (ALS)**

Sharyn Rosenberg, the Governing Board Assistant Director for Psychometrics, noted that the NAEP legislation specifies that the Governing Board is responsible for developing achievement levels for each subject area and grade tested by NAEP. In 1995, the Board adopted a policy on Developing Student Performance Levels for the National Assessment of Educational Progress; this policy is used to guide procurements on NAEP achievement levels setting. Ms. Rosenberg noted that COSDAM received an overview of the TEL ALS project at the August meeting, just a few weeks after the contract was awarded to NCS Pearson (Pearson). Ms. Rosenberg stated that a lot of work on the project has occurred during the last few months, and several changes have been made.

Paul Nichols, the TEL ALS project director at Pearson, provided an update on recent project activities. Since the last COSDAM meeting on August 1<sup>st</sup>, the Technical Advisory Committee on Standard Setting (TACSS) met to discuss drafts of the Planning Document and Design Document. Following recommendations of the TACSS, Pearson will present the results from the achievement levels setting study for Board action at the August 2015 meeting rather than the May 2015 meeting, as originally planned.

Mr. Nichols provided an overview of the Design Document and highlighted the following changes. Pearson will use a modified item mapping (i.e., Bookmark) methodology, and the entire standard setting process will be computerized. As a subcontractor to Pearson, Measurement Incorporated (MI) will provide the standard setting software, which is based on a modification of the software used for the recent Smarter Balanced Assessment Consortium (SBAC) standard setting activities. A usability study will be conducted in early December to provide feedback on how panelists navigate between the two computers. COSDAM members did not have any questions about the project changes or about the Design Document.

Mr. Nichols noted that the Design Document is currently out for public comment, from October 29<sup>th</sup> until November 28<sup>th</sup>, but no comments have been received thus far. Pearson has a web page for public comment, and the link has been sent to over 100 interested organizations as part of the panelist nomination effort; there is also a link to the public comment site from the Governing Board home page. Jim Geringer expressed concern about not receiving any comments; other COSDAM members noted that the document is very long and technical.

Mr. Nichols requested feedback from COSDAM on a proposed plan to collect public comment on the *outcomes* of the TEL ALS. The Governing Board policy on developing student performance levels specifies that public comment be sought at critical junctures throughout the process, including for the proposed levels. However, the proposed levels traditionally have been treated as embargoed data and have not been distributed prior to the Report Card release. Mr. Nichols proposed that a small group of interested stakeholders (such as state testing directors) be convened in late June 2015 in conjunction with the National Conference on Student Assessment (NCSA) in San Diego. Participants would be required to sign a confidentiality agreement, and the results of the ALS process would be shared. COSDAM members agreed that this proposal was reasonable but requested that participants include policymakers in addition to state testing directors. COSDAM members expressed interest in revisiting the 1995 Board policy on developing student performance levels but preferred to wait until after the report from the evaluation of NAEP achievement levels is released in 2016.

### **Update on Transition to Technology-Based Assessment (TBA)**

Andreas Oranje of Educational Testing Service (ETS) provided an overview of the current plans for transitioning NAEP assessments onto a digital platform with a focus on the draft white paper, which was sent to COSDAM members via email earlier that week. In 2015, the paper-based assessments in Reading, Mathematics, and Science will continue to be administered and used for reporting NAEP results. In addition, technology-based assessments will be administered as part of the TBA start-up process, for the purpose of conducting a mode study (examining potential differences in student performance attributable to the mode of administration) and exploring

whether and how the trend can be maintained. The 2015 TBA start-up activities will be based on existing paper-based items that have been “trans-adapted,” or transferred to a digital platform. In 2016, pilot tests will be conducted using new TBA items that do not have current paper-based equivalents, such as scenario-based tasks. In addition, the trans-adapted items will be re-administered in 2016 to determine the extent to which the new content can be scaled with the existing content. The current plan for 2017 is to administer the Reading and Mathematics assessments exclusively by tablets that would be provided by the NAEP administrators.

Mr. Oranje noted that the draft white paper was recently reviewed by three NAEP expert panels: the Design and Analysis Committee (DAC), the NAEP Validation Studies panel (NVS), and the Quality Assurance Technical Panel (QATP). Feedback from these expert groups will be incorporated in the final version of the white paper. The following suggestions were made by the expert panels: 1) disentangle TBA from MST (multi-stage testing), which is not inherent in the design; 2) address digital equity/fairness in the TBA studies; 3) add a state validation component, which is not currently planned and would involve significant costs; 4) provide an argument for the value of scenario-based tasks; and 5) add evaluation criteria for the decision about whether trend can be maintained. In particular, COSDAM members were interested in exploring a state validation component and retaining a paper-based component in 2017.

COSDAM members emphasized the critical importance of the 2017 NAEP results and maintenance of trend, given all of the changes occurring in state assessments. Terry Holliday stated that if the cost of moving to TBA is that we lose the trend, then NAEP’s gold standard will be undermined. There was consensus that everything possible must be done upfront to maintain the trend, and that the question should be reframed as *how* rather than *whether* trend can be maintained. There was considerable discussion about the extent to which the trend decision is a policy issue. It is unlikely that the data from the bridge studies will be definitive, and the narrative around the trend decision (including any caveats) will be as important as the trend decision itself.

### **Upcoming Procurement: Review of Existing Studies on Motivation and Engagement in NAEP**

Ms. Rosenberg noted that there have been several previous discussions about the tendency for people to question whether students (especially at grade 12) put forth their best effort on a low-stakes assessment. She described an upcoming procurement to conduct a critical review and summary of research on student motivation and engagement on NAEP, with the following goals: 1) to critically evaluate the claims that have been made; 2) to summarize the extent to which results are consistent across studies; and 3) to recommend future research that should be performed.

COSDAM members offered the following input on the procurement: include grades 4 and 8 in addition to grade 12; consider research related to motivation and engagement on state tests as well as NAEP; consider the extent to which findings from previous studies may still be current; consider effects of item types and mode of administration; and keep the budget for the procurement modest.

**Approaches to Innovation**

Finally, COSDAM members were asked to reflect on Chairman Terry Mazany’s charge to discuss ideas for innovation. Mr. Geringer asked, “Innovation for what purpose – are we asking the right question?”

The following topics were raised related to innovation: the move to TBA; exuberance in how we solve the problem of maintaining trend, not just the move to TBA itself; the process by which we develop items that measure instructional sensitivity; comparing state trends to NAEP trends; using gaming data to gather dynamic and realtime feedback to inform learning; and increasing NAEP’s web presence through data visualization.

I certify the accuracy of these minutes.

*Louis M. Fabrizio*

12-15-14

\_\_\_\_\_  
Lou Fabrizio, Chair

\_\_\_\_\_  
Date