

National Assessment Governing Board

Committee on Standards, Design and Methodology

Report of February 28, 2014

COSDAM Members: Chair Lou Fabrizio, Vice Chair Fielding Rolston, Lucille Davy, Andrew Ho, Terry Holliday, James Popham, and Leticia Van de Putte.

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

Other Attendees: John Easton, Director of the Institute of Education Sciences and ex officio member of the Governing Board. NCES: Janis Brown, Samantha Burg, Jing Chen, Andrew Kolstad, Drew Malizio, Taslima Rahman, and Bill Tirre. AIR: George Bohrnstedt, Young Yee Kim, and Gary Phillips. CCSSO: Katie Carroll. ETS: Longjuan Liang, Rochelle Michel, and Andreas Oranje. Hager Sharp: Lauren Werner. HumRRO: Laurie Wise. Optimal Solutions Group: Rukayat Akinbiyi. Pearson: Brad Thayer. Westat: Keith Rust.

Introductions and Review of Agenda

Lou Fabrizio, Chair of the Committee on Standards, Design and Methodology (COSDAM), called the meeting to order at 10:05 a.m. and welcomed members and guests. Mr. Fabrizio noted that the agenda included presentations on plans for two-stage adaptive testing in mathematics and science, results from the NAEP-PIRLS linking study, and a closed session on participation and engagement data for the 2013 grade 12 mathematics and reading assessments.

Plans for two-stage adaptive testing in NAEP Mathematics and Science

Sharyn Rosenberg, the Governing Board Assistant Director for Psychometrics, began by noting that the transition to technology-based assessments (TBA) is one of the greatest opportunities and challenges facing the NAEP program. The 2011 Writing assessment was the first fully technology-based NAEP assessment, followed by the 2014 Technology and Engineering Literacy (TEL) assessment. For the Reading, Mathematics, and Science assessments, work is currently underway to transition the existing paper-based assessments to a technology platform, based on the existing assessment frameworks. Ms. Rosenberg noted that this presentation is likely to be the first of many discussions about the TBA transition in the coming years. She introduced the presenter, Andreas Oranje.

Andreas Oranje of Educational Testing Service (ETS) began with an overview of the current plans for NAEP to transition to TBA in Reading, Mathematics, and Science by 2017. In 2015, the paper-based assessments will continue to be administered and used for the reporting of

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 transferred to a technology 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.

Mr. Oranje described the following reasons behind the intention to eventually implement some form of adaptive testing: increasing measurement precision across the distribution, describing performance across the distribution, and improving the student experience by targeting items to a given student's performance. ETS has been receiving input on the transition to TBA and the use of adaptive testing from several expert groups, including the Design and Analysis Committee (DAC) and an additional TBA expert panel (Gage Kingsbury of the Kingsbury Center at NWEA, Hua-Hua Chang of University of Illinois at Urbana-Champaign, and Tim Davey of Educational Testing Service). Design considerations include the administration model (adaptive testing), type of adaptive testing (multi-stage), number of stages (most likely two, based on results from the 2011 Math Computer Based Study at grade 8), and number of levels (not yet determined). Additional considerations include how to implement adaptive testing in the context of scenario based tasks (i.e., *should the assessment be adaptive within a scenario-based task, between scenario-based tasks, or not at all for scenario-based tasks?*).

COSDAM members asked questions related to the general transition to TBA. First, Terry Holliday asked whether it is possible for NAEP to utilize the assessment items or technology platforms from the assessment consortia, given that they were developed using federal funds. Mr. Fabrizio requested that NCES explore the feasibility of this request and report back to COSDAM at a future meeting. There was also a question by Leticia Van de Putte about whether gaming principles would be used in assessment design; Mr. Oranje responded by providing some information about ongoing work in this area outside of the NAEP program.

Related to the use of adaptive testing, a question was raised by Mr. Holliday about whether the reported results will conclude that eighth-grade students in some states are performing at the average level of fourth-grade students in the nation. Ms. Rosenberg clarified that item development for the adaptive tests will adhere to the NAEP frameworks; for example, the grade 8 Mathematics assessment will contain a range of items from easy to difficult, but the content of all items will be based on that particular framework.

Andrew Ho expressed caution over moving too quickly with the transition to TBA, given the primary goal of maintaining the trend. One alternative is to proceed more slowly by first transitioning to TBA, then subsequently introducing multi-stage testing.

Results from NAEP-PIRLS linking study

Mr. Fabrizio began by noting that the Board approved a resolution in November 2009 to plan for 2011 linking studies between NAEP and TIMSS (Trends in International Mathematics and

Science Study), and NAEP and PIRLS (Progress in International Reading Literacy Study). He noted that the Board had heard several presentations about the NAEP-TIMSS linking study, which was released last fall. He welcomed presenter Gary Phillips.

Gary Phillips of the American Institutes for Research (AIR) began by describing the similarities and differences between the 2011 NAEP-TIMSS linking study and the 2011 NAEP-PIRLS linking study. The 2011 NAEP-TIMSS linking study in grade 8 Mathematics and Science compared the use of three different methodologies: calibration, projection, and statistical moderation. The 2011 NAEP-PIRLS linking study in grade 4 Reading used only statistical moderation, a procedure which matches the NAEP and PIRLS means and standard deviations.

Mr. Phillips noted that the content and readability of NAEP and PIRLS is similar, but there are some differences between the two assessments in terms of sampling procedures, testing window, testing time, accommodations, and passage length. When comparing the NAEP standards (Basic, Proficient, Advanced) with the PIRLS standards (Low, Intermediate, High, Advanced), the study found that the NAEP benchmarks were higher than the corresponding PIRLS benchmarks. When expressed on the NAEP scale for grade 4 Reading, the NAEP Proficient achievement level corresponds to a score of 238, whereas the PIRLS High achievement level corresponds to a score of 217.

COSDAM members discussed the unintended implications of using the “Proficient” label in an educational landscape where “Proficient” is often interpreted to mean “on grade level.” The discussion included an acknowledgement that the NAEP achievement levels are still in a trial status, which may be resolved by the upcoming procurement to conduct an independent evaluation of the NAEP achievement levels. Mr. Holliday expressed a concern that the NAEP standards are too high compared to external benchmarks; Lucille Davy cautioned against chasing a low standard. Most of the discussion about the appropriateness of the NAEP standards was related to the “Proficient” label rather than the placement of the cut scores.

Mr. Phillips noted that the linking study was conducted at the national level only; the National Center for Education Statistics (NCES) decided not to estimate state PIRLS means or state PIRLS international benchmarks. Mr. Phillips explained the reason behind this decision, which was that the 2011 NAEP-TIMSS linking study illustrated the importance of the correlation between the two assessments for making accurate predictions. For the 2011 NAEP-PIRLS linking study, the correlation between NAEP and PIRLS was unknown because only one state, Florida, participated in both assessments at the state level. In Florida, the PIRLS-equivalent state mean (predicted from the state NAEP scores) was not significantly different from the actual PIRLS state mean, providing some validity evidence for the linking study results. Cornelia Orr, Executive Director of the Governing Board, noted that states could estimate their own PIRLS-equivalent state means using the linking equation from this study.

Mr. Fabrizio asked why the NAEP-PIRLS linking study did not offer states the opportunity to participate in state PIRLS, and when the next opportunities would occur for conducting linking studies with PIRLS and TIMSS. John Easton, Director of the Institute of Education Sciences and ex officio member of the Governing Board, responded that the budget could not support state participation in PIRLS. Ms. Rosenberg noted that the NAEP and PIRLS administration windows

are not scheduled to align until 2021, but that NAEP and TIMSS are both scheduled to be administered in 2015. Drew Malizio of NCES added that there are plans to conduct a 2015 linking study between NAEP and TIMSS, using statistical moderation only.

CLOSED SESSION 11:25 a.m. – 12:05 p.m.

COSDAM Members: Chair Lou Fabrizio, Vice Chair Fielding Rolston, Lucille Davy, Andrew Ho, Terry Holliday, James Popham, and Leticia Van de Putte.

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

Other Attendees: John Easton, Director of the Institute of Education Sciences and ex officio member of the Governing Board. NCES: Janis Brown, Samantha Burg, Andrew Kolstad, Drew Malizio, Taslima Rahman, and Bill Tirre. AIR: George Bohrnstedt and Young Yee Kim. ETS: Longjuan Liang, Rochelle Michel, and Andreas Oranje. Hager Sharp: Lauren Werner. HumRRO: Laurie Wise. Optimal Solutions Group: Rukayat Akinbiyi. Westat: Keith Rust.

In accordance with the provisions of exemption (9)(B) of Section 552b(c) of Title 5 U.S.C., the Committee on Standards, Design and Methodology met in closed session on February 28, 2014 from 11:25 a.m. – 12:05 p.m. in order to discuss information regarding analyses of the 2013 grade 12 Mathematics and Reading assessments, including secure data.

Participation and Engagement on the 2013 NAEP Grade 12 Mathematics and Reading Assessments

Ms. Rosenberg introduced the session by noting that COSDAM has previously discussed the tendency for some people to question whether grade 12 results represent students' best efforts. There is research on NAEP and motivation that has been done and is often quoted but has not been critiqued. As part of the 2009 release of the grade 12 results, NCES did present some evidence that grade 12 students do take the test seriously, such as participation rates and completion of open-ended questions; these data are now available for the 2013 grade 12 Mathematics and Reading assessments. She introduced presenter Samantha Burg.

In closed session, Samantha Burg of NCES presented data on grade 12 school and student participation rates and item response rates. The presentation included trend data for grade 12 (from 1992 through 2013) and comparisons to grades 4 and 8. Mr. Ho asked whether the information about participation and engagement will be included in the grade 12 report. Ms. Burg responded that a "Focus on NAEP" publication about grade 12 participation and engagement will be released in conjunction with the 2013 NAEP grade 12 Report Card.

The presentation included information about the strategies that schools use to encourage students to put forth their best efforts on NAEP, such as: thanking students for participating in NAEP, having students and teachers urge students to participate, and holding an assembly or pep rally for NAEP.

OPEN SESSION: 12:05 p.m. – 12:15 p.m.

(See list of attendees for the closed session. No additional attendees joined this session.)

Committee Questions on Information Items

In closing, Mr. Fabrizio asked whether there were questions on any of the information items: 1) NAEP 12th Grade Academic Preparedness Research: Phase 2 Research Updates; 2) Update on TEL Achievement Levels Setting Procurement; 3) Update on Implementation of SD/ELL Exclusion Policy; or 4) Update on Evaluation of NAEP Achievement Levels Procurement.

Fielding Rolston asked about the status of the TEL achievement levels setting (ALS) procurement. Ms. Rosenberg responded that the TEL ALS will take place in two stages via separate procurements. First, the achievement levels descriptions (ALDs) for grade 8 must be finalized. A statement of objectives was developed, and a Request for Quotations was issued in early February; the contract is expected to be awarded by the first week in April. The final achievement levels descriptions will be presented to COSDAM and the full Board for approval at the July 31 – August 2, 2014 meeting. The second procurement is for the 2014 TEL achievement levels setting at grade 8. A pre-solicitation notice was posted in mid-February, and the Request for Proposals is expected to be issued in March.

Ms. Van de Putte asked about the status of implementing the SD/ELL exclusion policy. Ms. Rosenberg noted that the COSDAM and Reporting and Dissemination Committees had met in joint session during the December 2013 Board meeting to discuss the 2013 participation data for grades 4 and 8 Reading and Mathematics. Contrary to Board policy, NCES has continued to permit schools to exclude students whose IEPs called for accommodations that NAEP does not allow. The staffs of NAGB and NCES are considering proposals for potential revisions to the policy that would allow the spirit of the policy to be implemented in a manner that NCES considers technically sound; the intention is to hold a follow-up joint session during the May Board meeting.

Ms. Van de Putte responded that many Spanish-speaking students are not ready to be tested in English, and that some students continue to be taught and assessed in Spanish for several years. Ms. Orr noted that the December 2013 joint session focused primarily on the implementation of the policy for students with disabilities, but that the policy implications for English language learners should also be discussed.

I certify the accuracy of these minutes.

Louis M. Fabrizio

Lou Fabrizio, Chair

March 6, 2014

Date