

National Assessment Governing Board

Committee on Standards, Design and Methodology

Report of December 2, 2011

COSDAM Attendees: Lou Fabrizio (Chair), Tonya Miles (Vice Chair), John Q. Easton (*Ex officio* member of the Governing Board and Director of the Institute of Education Sciences), Terry Holliday, Jim Popham, Andy Porter, Fielding Rolston, and Jennifer Ranji (representing Governor Jack Markell).

Governing Board Staff: Executive Director Cornelia Orr, Susan Loomis, Michelle Blair, and Ray Fields.

Other Attendees: NCES: Commissioner of Education Statistics Jack Buckley, Janis Brown, Jing Chen, Steve Gorman, Eunice Greer, Andrew Kolstad, Drew Malizio, Bill Tirre, and Ebony Walton. AIR: Fran Stancavage. Data Recognition Corporation and REDA International Inc: Herbert Baum. ETS: Dave Freund, Steven Lazer, John Mazzeo, and Rebecca Moran. Hager Sharp: Joanne Liu. HumRRO: Monica Gribben. Measured Progress: Luz Bay. MetaMetrics: Heather Koons and Malbert Smith. NAEP ESSI: Burhan Ogut and Sharyn Rosenberg. Pearson: Connie Smith. Virginia Department of Education and Chair of the CCSSO/NAGB Policy Task Force: Pat Wright. Westat: Keith Rust.

Lou Fabrizio, Chair of the Committee on Standards, Design and Methodology (COSDAM), called the meeting to order at 10:20 a.m. and welcomed members and guests. Mr. Fabrizio welcomed two new members to the Governing Board and to COSDAM: Fielding Rolston and Terry Holliday. Leticia Van de Putte was unable to attend, and Jennifer Ranji attended as Governor Markell's representative.

Mr. Fabrizio reminded everyone of the closed session beginning at noon and the need to stay on schedule when discussing the other topics on the agenda.

2009 Grade 12 NAEP Preparedness Research

Almost half of the meeting time was devoted to discussion of Grade 12 NAEP Preparedness Research results. Mr. Fabrizio referred attendees to Attachment A of the COSDAM briefing materials for the overview of the program of preparedness research adopted by the Board in March 2009.

1. **NAEP-SAT Linking Study:** Mr. Fabrizio noted that the first topic on the agenda was an informational item to present a written report describing the procedures and highlighting results of the analyses conducted by ETS for establishing the statistical relationships for grade 12 NAEP in both reading and mathematics with the grade 12 students having taken the SAT by June 2009. Susan Loomis reminded members that ETS staff have reported to COSDAM at several of the last meetings, and she noted that ETS researchers were on hand to answer any questions regarding the research on the statistical relationship for the national NAEP sample with the SAT.

There were no questions, but Andy Porter noted that the high correlation (.91) between NAEP and SAT math provides some indication that motivation of 12th graders for performance on NAEP does not appear to be the problem that many assume to be the case. He also noted the finding that the knowledge, skills, and abilities needed for job training in occupations is not the same as that for performance on NAEP seems entirely reasonable.

2. NAEP-Florida Data: Rebecca Moran of ETS presented results for the study of statistical relationships of NAEP with data from the Florida Department of Education (FLDOE) Data Warehouse. The results included information about average NAEP scores for Florida students who meet the SAT “college success” benchmark and those who do not. The SAT benchmark is 1550 for the composite score and 500 for each of the separate tests in mathematics, critical reading, and writing. The data for Florida students revealed the same general findings as for the NAEP-SAT national samples. The range of preparedness is near the Proficient cut score for grade 12 NAEP in both reading and mathematics.

Andy Porter requested clarification about the meaning of the NAEP data for Florida and the importance of the research findings. Ms. Moran noted that the Florida data are of limited value in that the data are for only one state; but, she noted that the data are very useful in providing the mutually confirmatory evidence sought for overall results. The overall pattern of findings based on the data from Florida reflected the statistical relationships based on the national sample for NAEP with the SAT. Terry Holliday noted that it is important for states to have the data on NAEP in relation to the SAT, ACT, and other assessments, and the FLDOE data represents a start in providing that information.

At the conclusion of Ms. Moran’s presentation, Mr. Porter suggested that the results perhaps show that performance at the NAEP Proficient achievement level by 12th graders is a good indicator of preparedness for college. This observation was consistent with the findings across the analyses of statistical relationships based on the national NAEP sample and on the Florida NAEP sample.

John Easton noted the difference between the average NAEP score for Florida students who attended a two-year versus a four-year college after high school. He wondered if it would be necessary to report preparedness for two-year and four-year colleges separately. Ms. Loomis noted that students in Florida who need remediation must take the remedial courses in two-year colleges; no remedial courses are offered in the four-year colleges and universities. This might tend to increase the difference in achievement of students entering the two types of institutions immediately after high school. Cornelia Orr further noted that the distinction between two-year and four-year colleges in Florida is blurred and differs from that in most other states. Some colleges that were formerly community colleges and continue to function primarily as a two-year college now offer a bachelor’s degree in some programs. Remedial courses are offered in that special category of Florida’s public four-year colleges.

3. Judgmental Standard Setting Studies: Susan Loomis provided an overview of the key findings from the judgmental standard setting studies. COSDAM meet on November 4, 2011 via a webinar for a comprehensive review of findings. The Committee received a report in advance of that meeting, and the December briefing materials included selected highlights of the findings.

The studies were designed to set a cut score to represent the minimal level of academic preparedness required for students to be placed in a credit-bearing college course of the sort that satisfies a general education requirement OR the minimal level of academic preparedness required for students to enter a job training program in one of five different occupations: automotive master technicians; computer support specialists; licensed practical nurses; heating, air conditioning and ventilation technicians; and pharmacy technicians. Replicate panels were used for each study to provide information about the reliability of the results. Ms. Loomis reported that the replicate panels produced similar results (replicated) in some cases, but certainly not for all. The panelists had difficulties with the tasks of describing the knowledge, skills, and abilities required for correctly answering the questions on NAEP. Panelists noted that many items seemed inappropriate for the requirements needed by students in their programs. A special study was conducted at the last operational session that collected information on the items panelists rated as “irrelevant.”

Additional research is underway to produce more information regarding the outcomes of the judgmental standard studies. Materials are being collected for courses taken by students entering the job training programs included in the studies: course syllabi, tests, reading assignments, and text books are being examined to compare to the NAEP items that represent “preparedness” at the cut scores and to compare to the descriptions used to set the cut scores for academic preparedness in the job training programs for each occupational area.

Jim Popham expressed appreciation for the openness and frankness with which staff discussed the findings of the judgmental standard setting studies and appreciation for the effort to collect more information to better understand the results of the studies. But, he also expressed concern about the use of an assessment developed for what he considered to be an entirely different purpose to measure preparedness. John Easton questioned whether a different assessment should be developed to measure preparedness. Andy Porter indicated that he still sees potential for reporting preparedness on NAEP—especially for college. He noted that Achieve had worked with the Board to modify the frameworks for measuring preparedness. Lou Fabrizio noted, however, that Achieve focuses on “high trajectory jobs” that are high-paying and likely require more training than the occupational areas the Governing Board has examined.

4. Validity Framework: In addition to reviewing findings from the preparedness research studies, COSDAM was asked to discuss what the findings mean and how they contribute to our understanding of preparedness. A validity framework is being developed by Board staff that will present statements of findings and the evidence from the various studies that will be offered in support of the statements.

Ms. Loomis reminded the Committee of the general approach adopted by the Board to conduct the broad array of studies to collect evidence to be evaluated for mutual confirmation. She noted that the validity framework has been presented to the Committee on a few previous occasions for Committee input on the organization, level of detail, and so forth. At this meeting, staff asked COSDAM to discuss what the results show and to recommend statements about preparedness findings the Board wants to make and feels are supported by the research findings.

Ray Fields reiterated these points and elaborated on ideas that COSDAM might consider for reporting.

- a.) A score of 150 on the NAEP math scale represents the score at which the 25th percentile for Florida students who needed no remediation in mathematics intersects with the score at the 75th percentile for Florida students who needed remediation in mathematics. He asked COSDAM to consider whether this might serve as an indicator of minimal preparedness.
- b.) Given the data linking NAEP and SAT performance, it would be possible to report on “prepared for success” for grade 12 NAEP. The statistical studies have produced a score on NAEP representing high probabilities (.50, .67, and .80) of scoring at the SAT benchmark that represents a freshman year grade-point average of B- or higher. “Prepared for freshman year success” seems a possibility for NAEP reporting.
- c.) As noted earlier by Mr. Porter, it also seems possible to interpret the Proficient achievement level for grade 12 in terms of preparedness. The NAEP proficient cut score is approximately the same as the reference point for the SAT college success benchmark score.

Jim Popham stated that he had carefully reviewed the validity framework. He noted that it is very important to distinguish the level of confidence or uncertainty associated with the study findings and not simply talk about “relevance” of NAEP for reporting preparedness.

Andy Porter suggested that staff consider using the standard setting method developed by former COSDAM Chair Ed Haertel and implemented by Pearson for the Achieve algebra II tests. That procedure includes having panelists not only set cut scores, but also review additional evidence and recommend an overall cut score based on the entire compilation of evidence. Ms. Loomis noted that this procedure is used for reviewing NAEP achievement levels results, except the review of additional evidence is by COSDAM and the Governing Board, rather than the standard setting panels. She also noted that the current plan for vetting the findings of the NAEP grade 12 preparedness research includes a comprehensive review of findings across all studies by technical advisors, content experts, and representatives of NAEP stakeholder groups, as well as by COSDAM and the Governing Board.

Mr. Porter noted that a vetting process was not likely to produce the same outcome as the standard setting procedure he was recommending. Mr. Fabrizio stated that COSDAM needs a clear recommendation based on a determination of statements and propositions about preparedness on NAEP that are supported by the evidence generated by the studies. COSDAM asked staff to develop a plan for producing a clear recommendation and report back to the Committee in March 2012.

Terry Holliday asked about how NAEP will report on preparedness if preparedness for NAEP is not the same as for states. He noted that states should be given advance notice if there is the possibility of having NAEP report that only half as many students are “prepared” as the number reported by other studies—possibly including the results for the Common Core State Standards when the two consortia’s assessments are administered.

Making a Difference

Each standing committee was asked to schedule 20-30 minutes for discussion of the proposed actions the Governing Board might take to “Make a Difference” in raising student achievement and closing achievement gaps. COSDAM had very little time to discuss this topic since a closed session was scheduled to begin only about 10 minutes later. Mr. Fabrizio suggested that COSDAM members write their ratings and report them to him. Tonya Miles urged COSDAM colleagues to think broadly and not be focused on concerns about the limitations to Governing Board authority because staff could sort through what is/is not allowed for Board action.

Jim Popham objected to the activity on the grounds that the charge needed to be clarified before requesting members to rate preferences and priorities for reporting. He noted that the statement of the goal in the “Making a Difference” tab was not the same as that in the summary for COSDAM and the other committees. He noted the difference between a goal to close achievement gaps and a goal to raise awareness of the need to close achievement gaps. Ms. Miles said that she agreed in some respects, but she could also see one as a subset of the other. She mentioned that there are some things that the Board cannot do. But, with the data that are available, she felt that the Board has a responsibility to share the data with others and to encourage them to use the data to make a difference.

Mr. Fabrizio reiterated his request for COSDAM members to give him their ratings before the start of the morning session on December 3, 2012. Mr. Popham again objected and expressed his feeling that this important issue was being treated as a triviality. Mr. Porter asked to have Mr. Popham’s objections noted in the report by Mr. Fabrizio to the Board, and Mr. Fabrizio assured everyone that this was his intention, and that it was also his intention to carry out the request of Chairman Driscoll to have the Committee provide their ratings of the proposals.

John Easton recommended adding Kati Haycock’s suggestion from her Friday morning presentation to the Governing Board to have NAEP provide support to states in preparation for having results from the Common Core State Standards, and his recommendation was generally accepted by the members. Mr. Fabrizio reviewed the 10 proposals to be rated by each COSDAM member.

The members submitted their ratings, the results were tallied, and Mr. Fabrizio presented the COSDAM ratings to the Governing Board for discussion in a general session on Saturday, December 3, 2011.

CLOSED SESSION 12:00 - 12:30 p.m.

Achievement Levels for 2011 Writing NAEP at Grades 8 and 12

COSDAM Attendees: Lou Fabrizio (Chair), Tonya Miles (Vice Chair), John Q. Easton (*Ex officio* member of the Governing Board and Director of the Institute of Education Sciences), Terry Holliday, Jim Popham, Andy Porter, Fielding Rolston, and Jennifer Ranji (representing Governor Markell).

Governing Board Staff: Executive Director Cornelia Orr and Susan Loomis.

Other Attendees: NCES: Jing Chen, Steve Gorman, Eunice Greer, Andrew Kolstad, Drew Malizio, and Bill Tirre. ETS: John Mazzeo and Rebecca Moran. Hager Sharp: Joanne Liu. HumRRO: Monica Gribben. Measured Progress: Luz Bay. Pearson: Connie Smith.

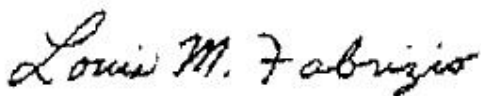
In accordance with the provisions of exemption (9)(B) of Section 552b(c) of Title 5 U.S.C., the Committee on Standards, Design and Methodology met in closed session on December 2, 2011 from 12:00 p.m. to 12:30 p.m. in order to review and discuss reports including secure data and results of research conducted to set achievement levels cut scores for the National Assessment of Educational Progress in writing.

The Committee was briefed on the field trial and pilot study for the writing achievement levels-setting process by Luz Bay, Assistant Vice President for Client Services at Measured Progress and Project Director for the NAEP Writing Achievement Levels-Setting Process. She provided recommendations developed from the field trial that were incorporated in the design of the process implemented for the pilot study. The goal of the pilot study was to implement the procedure exactly as planned for the operational achievement levels setting (ALS) process scheduled for February 2012. Ms. Bay provided information on the computerized version of the Body of Work (BoW) standard setting process developed for the NAEP writing ALS and called BoWTIE (Body of Work Technological Integration and Enhancement). She showed results of the pilot study and results of a special study conducted at the end of the pilot study. For the special study, panelists used the achievement levels descriptions developed for the 2011 writing NAEP with the BoW methodology for classifying Bodies of Work (student test booklets) from the 2007 writing NAEP in grades 8 and 12.

COSDAM supported Ms. Loomis' recommendation to conduct an additional review of the writing achievement level descriptions by content experts prior to the operational study. If that review results in recommended changes to the descriptions, Mr. Popham recommended that a small-scale panel study be conducted using the revised descriptions prior to the operational achievement levels-setting process scheduled for February 7-10, 2012. The Board is scheduled to take action on the 2011 Writing NAEP achievement levels at its May 2012 meeting.

The COSDAM meeting was re-opened at 12:30 p.m. Mr. Fabrizio then adjourned the meeting.

I certify the accuracy of this report.



Louis M. Fabrizio, Chair

12-8-11

Date