

National Assessment Governing Board

Committee on Standards, Design and Methodology

March 2, 2012

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, Andy Porter, Fielding Rolston, and Leticia Van de Putte.

Governing Board Staff: Susan Loomis and Ray Fields.

Other Attendees: NCES: Andrew Kolstad. AIR: George Bohrnstedt. ETS: Steven Lazer and Andreas Oranje. HumRRO: Laress Wise. Measured Progress: Luz Bay. MetaMetrics: Eleanor E. Sanford-Moore. Westat: Nancy Caldwell and Marcie Hickman. Widmeyer: Jason Smith.

Lou Fabrizio, Chair of the Committee on Standards, Design and Methodology (COSDAM), called the meeting to order at 9:50 a.m. and welcomed members and guests. Mr. Fabrizio noted that neither Governor Jack Markell nor Jim Popham would be able to attend this meeting of COSDAM.

Mr. Fabrizio noted that there were only two topics on the agenda for this meeting, and both were important topics requiring considerable discussion. Mr. Fabrizio reminded everyone of the closed session on the achievement levels for the 2011 writing NAEP beginning at 11:30 a.m.

2009 Grade 12 NAEP Preparedness Research

Over half of the meeting time was devoted to discussion of Grade 12 NAEP Preparedness Research results. Mr. Fabrizio noted that there would be a presentation to the full Board on Saturday morning to share information with the members that COSDAM has been reviewing and discussing for the preparedness research project. He then asked Susan Loomis to introduce the Committee members to the topic and direct members to documents for discussion.

Ms. Loomis mentioned the lengthy period of study on the issue of 12th grade preparedness by the Governing Board. She noted that Andy Porter, finishing his second term on the Board, is the most tenured member of COSDAM. Prior to his membership on the Governing Board, Mr. Porter wrote a report for the Board on college preparedness in 2004. She also noted that John Easton was a member of COSDAM when the Board first became engaged in the issue of reporting 12th grade preparedness. Now, the first cycle of research for the 2009 NAEP is ending and staff are preparing to report results, planned for the summer/fall of 2012.

1. NAEP-Florida Study: Ms. Loomis 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 sample of Florida students who were in grade 12 in 2009. Sufficient numbers of the Florida grade 12 sample had scores on the ACT, SAT, and ACCUPLACER tests to link their scores in mathematics and in reading to the NAEP assessments in these two subjects. In addition, the Governing Board staff arranged an agreement with the

Florida Department of Education (FLDOE) to provide data on students in their post-secondary years for this preparedness research project. The analyses of data for Florida in relation to performance on NAEP have been presented to COSDAM at several previous meetings, and the report prepared for the March 2012 COSDAM meeting was a final report including all of the information to be reported for Florida performances on the 2009 NAEP.

ETS staff have reported to COSDAM at several previous meetings, and ETS researchers were on hand to answer any questions regarding the research on the statistical relationship for the NAEP state representative sample of Florida 12th graders and their performance on other tests and their performance in their first year of college. The freshman year post-secondary data are only for students attending a public institution in the state of Florida. Governing Board staff still expect to get data for students attending private and out-of-state colleges. Ms. Loomis stated that Florida had just recently provided data for the second post-secondary year, and those analyses will begin in the near future.

Mr. Fabrizio clarified that there were no new analyses in the report; the purpose of this session was to provide a final opportunity for COSDAM to ask questions or recommend further analysis of the Florida data. There were no questions specifically regarding the statistical analyses for Florida data and NAEP performance; however, Terry Holliday had a more general question regarding the purpose of the statistical relationship studies for NAEP with Florida data and how those results would be useful. His questions helped to focus discussion on how the results would be useful to states and why states would want to participate in 12th grade NAEP. Discussion of these questions helped to clarify the purpose for the preparedness research, in general.

2. Judgmental Standard Setting Studies: Susan Loomis summarized the findings of the judgmental standard setting studies and the conclusions of the Governing Board staff regarding the results of the studies. She noted that the staff recommendation is that reference points or cut scores on the NAEP scale not be reported for these studies. Staff reached this conclusion because the results do not appear to be reasonable and are not consistent with other data sources.

Andy Porter noted that the finding that the current NAEP assessments are not well aligned to the requirements of the job training programs seems entirely likely. He noted that a good assessment of college preparedness would not necessarily be a good assessment of job training preparedness. Mr. Easton questioned whether the assessment developed for the Program for International Student Assessment (PISA) would likely be a better measure of student preparedness for job training programs than NAEP.

Additional research is underway to produce more information regarding the outcomes of the judgmental standard studies. WestEd and the Educational Policy Improvement Center (EPIC) are conducting this research for the Governing Board to delineate the knowledge, skills, and abilities in mathematics and reading required for students entering these job training programs and how these compare to the knowledge, skills, and abilities (KSAs) assessed by grade 12 NAEP. Materials are being collected for courses taken by students entering the job training programs included in the studies—approximately 20 job training programs for each of the five occupational areas and for each subject: course syllabi, tests, assignments during the early part of the course, and text books are being examined to compare to the NAEP assessment objectives, items that represent “preparedness” at the cut scores, and the descriptions used to set the cut scores for academic preparedness in the job training programs for each occupational area. The same types

of materials will be collected from a course taken at completion of the job training program to develop descriptions of the mathematics and reading skills required by students completing the programs. Ms. Loomis shared data from only two job training programs in the automotive master technician occupational area for which the first set of analysis has been completed. Based on comparisons of course components for those two programs relative to 12th grade NAEP framework components, about three-fourths of the KSAs included in the NAEP objectives for the grade 12 mathematics assessment are absent from the course materials for the two programs. For reading, however, slightly over half of the 12th grade framework objectives are represented by KSAs included in the beginning courses for these two programs.

In response to Terry Holliday's question about why Kentucky would want to participate in grade 12 NAEP when they already have data on ACT and WorkKeys performance for their students, Andy Porter responded that NAEP is a nationally representative test that provides the opportunity to compare performance of KY students to that of students in other states. The ACT data do not provide this. Andy also noted that having more than one definition of preparedness is helpful, so KY would benefit from data about both ACT and NAEP benchmarks.

Mr. Fabrizio reminded COSDAM that all of the results across all the studies need to be considered together.

3. Survey of Placement Tests and Scores in Higher Education: Ray Fields noted that the results of the survey are still being verified, but he did provide some details on the response rates by types of institutions, in addition to the 86% overall response rate. He reported that 91% of the 4-year public institutions in the sample responded, and 89% of the 2-year public institutions responded. Rates were a bit lower for private 4-year institutions (86%) and 2-year private institutions (76%). Final results will be available for consideration by the Technical Advisors for 12th Grade Preparedness Research at the April 5, 2012 meeting, and by COSDAM in May 2012.

4. Reporting Results of Preparedness Research for 2009 NAEP: Ms. Loomis reviewed plans for reporting results in two separate reports—a “public report” of about 40 pages directed to a general audience and giving the overall findings; and a “technical report” including far more detail and linking readers to source documents reporting on each of the research. Both reports will be available as on-line documents and the public report will be available in print.

Governing Board staff will produce a report in the form of a validity framework including statements of findings and the evidence in support of those findings ranging from the general to the specific findings. This document will be presented to the Technical Advisors for Grade 12 NAEP Preparedness Research in their April 5, 2012 meeting and to other NAEP technical advisory groups and others with interest in NAEP preparedness research. These reviews will be collected for reporting to COSDAM in May 2012.

Mr. Fields raised the issue of consequential validity and asked how the fact that there are no individual student scores involved in our reporting might impact the decisions on what to report. This led to more discussion about the findings and how to report them.

In response to a hypothetical statement by John Easton that the Board would ultimately be setting policy on preparedness, Mr. Fields noted that this research is not necessarily aimed at setting a

standard for academic preparedness. The results reported may not be expressed in the form of a “standard.” Mr. Easton suggested that this research would be used as a resource for Board policy.

Mr. Easton cautioned that the advice of the Technical Advisors should not be considered as “conclusive” because the Board needs to consider recommendations from various sources. Committee members agreed that this is an important point for staff to consider when interacting with the advisory groups.

Ms. Loomis and Mr. Fields will work to frame the preparedness statements and evidence in support of those statements and to have them vetted so the results of this process are ready to report to COSDAM in May 2012.

CLOSED SESSION 11:30 a.m. – 12:15 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, Andy Porter, Fielding Rolston, and Leticia Van de Putte.

Governing Board Staff: Susan Loomis.

Other Attendees: NCES: Andrew Kolstad. AIR: George Bohrnstedt. ETS: Steve Lazer. Measured Progress: Luz Bay.

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 March 2, 2012 from 11:30 a.m. to 12:15 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 second field trial and the operational achievement levels-setting (ALS) study for the 2011 NAEP in writing by Luz Bay, Assistant Vice President for Client Services at Measured Progress and Project Director for the NAEP Writing Achievement Levels-Setting Process. Ms. Loomis, working with writing experts, had modified the achievement levels descriptions as a result of the findings from the pilot study and the comments of panelists in that study. The revised descriptions were tried out in a field trial conducted at Measured Progress offices in Dover, New Hampshire. Panelists for both grades 8 and 12 were recruited for the study—approximately 20 for each grade. Panelists were recruited from among the consultants who score assessments for Measured Progress. This role apparently impacted their judgments in this very scaled-back process lasting only one day, rather than four days. She noted additional modifications to the process that were implemented in the field trial to provide panelists more experience with student performances and scoring rubrics in the training stages of the process. The achievement levels descriptions were further modified, as a result of feedback from the panelists in the second field trial.

Apparently, the additional work on the achievement levels descriptions was effective, because panelists were generally pleased with the achievement levels descriptions. The results of the ALS study and a special study were reported to the Technical Advisory Committee on Standard Setting February 23-24, 2012, and they found the results to be well documented, technically sound, and reasonable with respect to other data on writing achievement levels.

Ms. Bay noted that the panelists for NAEP achievement levels-setting are always selected because of their outstanding accomplishments. For example, teachers have typically received awards for their teaching excellence. Ms. Bay stated that the panelists for the 2011 writing achievement levels-setting process were especially outstanding. She provided information on the publications by several panelists and noted other accomplishments of the outstanding panelists.

A study was conducted at the end of the achievement levels-setting panel meeting to provide a basis of comparing performance on the 2007 writing NAEP to the 2011 writing NAEP. For the study, 2011 achievement levels descriptions were used as the criteria for classifying student performance on the 2007 assessment into the levels of achievement.

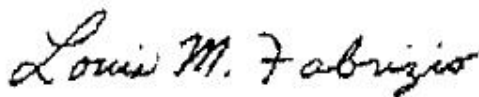
COSDAM was asked to consider how to represent student performance relative to the achievement levels. Ms. Loomis explained that the Board has considered selection of exemplar performances to be one of the major outcomes of the achievement levels-setting process. Panelists are asked to make recommendations regarding specific examples of student responses to represent performance at each level. The Technical Advisory Committee for Standard Setting had recommended criteria for selecting the examples, and selections were presented to COSDAM.

Luz Bay had prepared printed copies of the recommended exemplar performances. The recommendation is to represent performance at each achievement level with one student booklet judged by panelists to represent that level. This would result in one “booklet” of writing responses (2 tasks each) for each achievement level and grade—a total of 6 booklets (12 responses). Staff of the National Center for Education Statistics have expressed concern that this will require too much print space to report the exemplar performances. After considerable discussion, the conclusion was that the details of presenting this performance in the Nation’s Report Card can be resolved. COSDAM agreed that this is the preferred way to report performance relative to the achievement levels for Writing 2011 because it is consistent with the process for setting the writing achievement levels.

The Board is scheduled to take action on the 2011 Writing NAEP achievement levels at the May 2012 meeting in San Antonio.

The COSDAM meeting was opened at 12:15 p.m. at which time Mr. Fabrizio adjourned the meeting.

I certify the accuracy of this report.



Lou Fabrizio, Chair

March 12, 2012

Date