

National Assessment Governing Board

Committee on Standards, Design and Methodology

August 3, 2012

COSDAM Attendees: Lou Fabrizio (Chair), John Q. Easton (*Ex officio* member of the Governing Board and Director of the Institute of Education Sciences), Terry Holliday, Jim Popham, Andrew Porter, Leticia Van de Putte, and Fielding Rolston.

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

Other Attendees: NCEES: Associate Commissioner Peggy Carr, Gina Broxterman, Samantha Burg, Andrew Kolstad, and Taslima Rahman. AIR: George Bohrnstedt and Sharyn Rosenberg. CRP: Veleka Allen. Education Week: Catherine Gewertz. ETS: Steve Lazer and Andreas Oranje. Hager Sharp: Melissa Spade. HumRRO: Laress Wise. McGraw-Hill Education: Larry Snowwhite. MetaMetrics: Heather Koons. Optimal Solutions: Mark Partridge. Pearson: Connie Smith. Westat: Nancy Caldwell. Widmeyer: Nebyat Ejigu.

Lou Fabrizio, Chair of the Committee on Standards, Design and Methodology (COSDAM), called the meeting to order at 10:00 a.m. and welcomed members and guests. He noted that all members of the Committee were present at this meeting, with the exception of Governor Markell who has now resigned from the Governing Board.

Reporting Research for 12th Grade NAEP Preparedness Research: Validity Evidence for Reporting Preparedness on Reading and Mathematics NAEP

Mr. Fabrizio noted that COSDAM has monitored the research for 12th grade preparedness for almost a decade now. After a few observations regarding the program of research, Mr. Fabrizio asked Governing Board Staff Susan Loomis and Ray Fields to lead the COSDAM discussion of the 12th Grade Preparedness Reporting session. Susan Loomis provided a brief overview of the goals of the staff to be responsive to comments made by COSDAM and the full Board at the May 2012 meeting. The discussion at this August 2012 meeting should focus on the discussion questions included in the briefing materials distributed to the Board last week by Cornelia Orr. Mr. Fields noted that this report includes only a few sections of the report to let the Board see what we might want to say about the research. This partial report is intended to provide enough information to allow discussion of the questions posed to the Board.

The first discussion topic was audience. Staff attempted to shift the tone of the information reported in May and address the report more directly to a policy audience. Terry Holiday stated that the big question is whether the report provides convincing evidence to the public that “proficient” is now equated to college and career preparedness/readiness. He noted that Kentucky has tied the state proficient level to the ACT, Inc. college and career readiness standards. In Kentucky, Proficient is equivalent to college and career readiness. The Department of Defense equated ASVAB to the ACT for Kentucky. He emphasized that the audience for the report has to be state policy makers.

Andy Porter asked about the equivalence of college and career readiness. He noted that if the audience is largely states, then the Governing Board needs to know what states need and try to tailor our research and reporting to support their efforts. Otherwise, it will be difficult to be sure the state policy makers understand what the results mean.

Leticia Van de Putte noted that the business community in Texas wants results that are useful: What does it mean? She understands that part of the work of the Governing Board is to provide a rationale for why this preparedness research really matters. She notes that we have an unparalleled opportunity to show that it *does* matter. First, we have to know that we have it right—that the results are accurate. Then we need to show why it matters. She noted that the business community wants a workforce that can be trained—that has critical thinking skills. The importance of preparedness is not only with respect to having a qualified workforce, however; it is also important for producing consumers of products. Wage earners are needed for consumers of goods and services that businesses want to produce. She notes that if states can show that they have high levels of 12th grade students who are prepared for college and job training, they will attract new businesses. NAEP results can have an important impact, even without individual student preparedness scores to report. Having aggregate data for the states about the proportion of students that are prepared for college and job training and the level of preparedness they exhibit is a powerful factor for the future economy of states. The “preparedness” attribute is a stronger attraction than tax breaks as an influencing factor on locational decisions of businesses.

Mr. Holiday noted that the report needs to play-up the point that preparedness for college and career are *not* the same. This is an important finding, and it is contrary to what some others are reporting. It is important to get this message out to states so they do not just accept this and expect to find that prepared for college and career are the same. Kentucky has found that preparedness for college and career are not at the same performance level.

Mr. Fabrizio observed that there is currently no mention in the report of how the NAEP definition of preparedness and the research for NAEP differs from other studies. This is important to avoid misinterpretation of the NAEP results relative to the findings of other studies.

Mr. Fields asked COSDAM to comment on the tone of the report. Does it seem objective?

Tone was not a concern for Jim Popham, he was not sure that tone really matters. What matters, he said, is that the test results be actionable. We need to be sure that we get people involved in reviewing the results to answer whether the research findings matter. *Can* we answer the question of what the results mean and how the results matter?

Mr. Holiday said that he thinks the answer to *why* the results matter is that state leaders need data to show how their levels—proficient and college-career readiness levels—compare to the national level and to one another. States need some way to compare their results. Having the NAEP preparedness data will give states more information for judging whether their new college and career standards are reasonable. The new standard that states will have to set for proficient with the Common Core State Standards will be much higher than in the past, and the NAEP preparedness data will help states to see that the new levels are more in line with objective performance requirements.

Ms. Van de Putte said that we need to make states want to want to do this. We have to show them the data and then they will figure out how to use it. We can only give them evidence—we cannot make them use it. States have to be made to understand the data, and they will find how to use it.

Trial Urban District Assessment Policy Modifications (ACTION ITEM)

These changes clarify eligibility requirements and bring the policy in line with current practice. COSDAM reviewed these modifications in May 2012. Mr. Fields noted that a few very minor modifications were made after the May 2012 meeting, and those were highlighted in yellow.

Two districts (one for Texas and one for North Carolina) would be added to the list of eligible districts as a result of the proposed modifications. Lou Fabrizio and Leticia Van de Putte recused themselves for the vote on this item, and Tonya Miles chaired the committee discussion.

Mr. Porter recommended that the name of the central city/metropolitan region be included in the list of eligible districts, along with the school district name, so that identification of the “urban district” would be more evident.

Discussion clarified that this is a list of eligible districts, and not a list of districts that will actually participate in TUDA. The discussion also revealed the potential need for a policy to prioritize district participation in the event that funding falls so short that all eligible districts that wish to participate cannot be included.

Ms. Miles noted the need for deletion of one duplicate word on page 2.

Mr. Porter moved, and Fielding Rolston seconded, the motion for approval of the TUDA policy modifications. COSDAM unanimously approved the modifications, and the record shows that Ms. Van de Putte and Mr. Fabrizio were present, but not voting on this action item.

ACTION: COSDAM recommends approval by the Governing Board of the modifications to the TUDA policy, as appended in Attachment A to this report.

Linking NAEP and PISA

The Governing Board had urged the National Center for Education Statistics to develop a research plan for linking NAEP with the Program for International Student Assessment (PISA). Taslima Rahman of NCEES presented details of the study design for linking NAEP and PISA. The technical challenge of linking NAEP and PISA is rather large due to the differences in assessment programs: when in the school year the assessments are administered, the amount of time for testing, the reporting scale, accommodations (not offered in PISA), and achievement/performance levels. The study design was especially impacted by the fact that PISA assesses students in age groups (15-year olds will be in this study) and NAEP assesses students by grade (both students in grades 8 and 12 will be used in this study). The study will be for mathematics. Both a national sample and state samples will be included, and some students in grades 9, 10, and 11 will be sampled, in addition to the larger NAEP samples for grades 8 and 12. The design includes administration of grade 8 NAEP only, grade 12 NAEP only, and a “braided” booklet including both grades 8 and 12.

Mr. Popham expressed concerns about the need to explain what the results of the linking study mean. Mr. Porter followed-up that point by asking for an example of statement that could be made as a result of the study that cannot now be made. Mr. Porter anticipated that extensive “corrections” and caveats will be needed to explain the meaning of the study results and provide

for appropriate interpretations of findings. Peggy Carr noted that there will also be concern regarding the likelihood of large error bands around the state and national comparisons.

CLOSED SESSION 11:25 a.m. –11:50 a.m.

COSDAM Attendees: Lou Fabrizio (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, Leticia Van de Putte, and Fielding Rolston.

Governing Board Staff: Susan Loomis and Michelle Blair

Other Attendees: NCES Associate Commissioner Peggy Carr.

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 August 3, 2012 from 11:25 a.m. to 11:50 a.m. in order to discuss plans by NCES to issue a procurement for the purpose of evaluating NAEP achievement levels.

Independent Evaluation of NAEP Achievement Levels

Peggy Carr presented plans to COSDAM for an evaluation of NAEP achievement levels. Note that only NCES staff were allowed to participate with COSDAM and Governing Board staff for this closed session because it involves details regarding a procurement that has already been announced.

At the beginning of his appointment, NCES Commissioner Buckley announced his intention to have this evaluation to determine if the "warning" about the trial basis of achievement levels could be removed from the NAEP Report Cards. Ms. Carr noted that the plan is to issue the procurement for open and competitive bidding, rather than as a sole source procurement. The Department of Education's contracts office must make the final determination regarding this, however.

Ms. Carr noted that the reliability and validity of achievement levels would be the focus of the evaluation. Mr. Porter asked for an example of external validity that could be collected for NAEP achievement levels. Mr. Popham urged that "validity" be replaced by "defensibility."

Mr. Porter noted that it will be very difficult to collect information to evaluate achievement levels and very hard to evaluate the information. He suggested that NCES re-think the scope of work for this because he did not feel that the current plan would successfully accomplish the purpose of the evaluation.

CLOSED SESSION 11:50 a.m. – 12:30 p.m.

COSDAM Attendees: Lou Fabrizio (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, Leticia Van de Putte, and Fielding Rolston.

Governing Board Staff: Susan Loomis

Other Attendees: NCES: Samantha Burg, Andrew Kolstad and Taslima Rahman. AIR: George Bohrnstedt and Sharyn Rosenberg. ETS: Steve Lazer and Andreas Oranje. Hager Sharp: Melissa Spade. HumRRO: Laurie Wise. Pearson: Connie Smith. Westat: Nancy Caldwell.

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 August 3, 2012 from 11:50 a.m. to 12:30 p.m. in order to discuss information regarding a report including secure data and results of research conducted to explore using student response time in a computer based assessment as an indicator of student engagement.

Multi-Stage Adaptive Field Trial: Response Time Indicators of Engagement

The final presentation was by Andreas Oranje about the research for computer based NAEP assessments to determine if reliable estimates of student engagement can be captured from data on response time. Some information about this research strategy was presented to COSDAM and the Governing Board last year by Steven Wise, and his presentation showed the potential for how the computer based assessments in NAEP can be used to provide indications of student engagement.

The multi-stage adaptive field trial study results show little indication of “disengagement” by the 8th graders in the study. Additional research with 12th grade NAEP will be helpful.

Additionally, this methodology helps to more accurately sort “missing data” for coding as “omitted” responses and “not reached” items. Overall, the results showed an increase in the estimate of overall test difficulty, but there was no impact on the overall performance score on the assessment.

Mr. Popham described this as a “tremendous study.” He had praise for the great research design and implementation and for the fact that this research was a common sense approach to answer an important question.

Mr. Porter noted that this model is a very constrained model and that we can anticipate even more exciting measurement advances as NAEP pushes forward with research in computer adaptive testing.

Tonya Miles expressed some concerns about the use of computer adaptive testing in NAEP, especially about the response time assumptions about student behaviors.

Mr. Porter noted that more precise measurement of student ability, especially the lowest ability levels, is advanced with the use of adaptive testing. This benefits NAEP assessments in Puerto Rico, for example.

In closing, Mr. Fabrizio noted that this is the last meeting of COSDAM for both Andy Porter and Susan Loomis. He thanked Mr. Porter and Ms. Loomis for their service to COSDAM and the Governing Board.

The August 2012 meeting of COSDAM was adjourned at 12:30 PM.

I certify the accuracy of this report.

Louis M. Fabrizio

Lou Fabrizio, Chair

August 20, 2012

Date

Adopted: March 3, 2007

Revised: August 4, 2012



National Assessment Governing Board

Eligibility Criteria and Procedures for Selecting Districts for Participation in the National Assessment of Educational Progress

Trial Urban District Assessment

Policy Statement

Purpose

To define the eligibility criteria and selection procedures for participation of urban school districts in the National Assessment of Educational Progress (NAEP) Trial Urban District Assessment (TUDA).

Guiding Principles

Principle 1

Participation in TUDA shall be voluntary.

Principle 2

A primary goal of TUDA is to support the improvement of student achievement in the nation's large urban school districts and to focus attention on the specific challenges and accomplishments associated with urban education.

Principle 3

Districts participating in TUDA shall have the characteristics of large urban areas.

Principle 4

All districts that have participated in TUDA without interruption once included shall be deemed eligible and permitted to continue to participate.

Principle 5

The eligibility criteria for participation in TUDA shall promote (1) inter-district comparability, so that participating districts are reasonably similar with respect to key demographics and (2) efficiency in resources required of the NAEP program.

Principle 6

Increasing the total number of districts participating in TUDA shall be contingent on additional funding from Congress.

Principle 7

The Governing Board implements the selection procedures used to consider districts for participation in TUDA.

Principle 8

Districts applying for participation in TUDA should be committed to long-term participation.

Eligibility Criteria

1. Only cities having 250,000 or more population shall be represented in TUDA.
2. Districts participating in TUDA shall have a student enrollment large enough to support NAEP assessments in three subjects in each grade assessed. The enrollment requirement is a minimum of approximately 1,500 students per subject per grade level assessed.
3. Districts participating in TUDA shall have an enrollment district-wide or in the grade levels assessed that meets at least one of the following criteria:
 - a. 50% or more are minority students (i.e., African American, American Indian/Alaskan Native, Asian, Hispanic, Native Hawaiian/Other Pacific Islander, and/or multi-racial).
 - b. 50% or more are eligible for participation in the free and reduced-price lunch program (or other appropriate indicator of poverty status).

Districts that are very near to meeting a particular eligibility requirement may be considered eligible if they request to participate in the program and if funds are sufficient to permit participation. Eligibility data shall be updated and verified periodically.

Application and Selection Process/Procedures

To provide time for consultation, notification, and operational planning for the conduct of the Trial Urban District Assessments, the steps described below should be sequenced to conclude approximately 14 months prior to the start of testing.

1. Prior to the assessment year in which TUDA is to be conducted, the Governing Board Executive Director, in consultation with the National Center for Education Statistics (NCES), prepares a list of eligible districts and posts that list on the Governing Board website.
2. Prior to the assessment year in which TUDA is to be conducted, the Governing Board Executive Director sends a letter to each district that participated in the immediately preceding administration of TUDA to determine the district's interest in continuing as a participant in the upcoming administration of TUDA.
3. Based on funding from Congress and the decision of any previous TUDA participant not to continue, the Governing Board determines whether new districts can be considered for participation in the upcoming TUDA administration.
4. If the Governing Board determines that new districts can be considered for participation in the upcoming TUDA administration, the Governing Board Executive Director sends a letter notifying eligible districts of the opportunity to submit an application and the instructions for applying.
5. Eligible districts seeking to participate in TUDA submit an application to the Executive Director of the Governing Board. The application should be signed by the district superintendent or designee, include the most recent information documenting the district's enrollment and eligibility, and contain a commitment for long-term participation in TUDA if selected.
6. The Executive Director of the Governing Board and appropriate staff of the Governing Board shall review applications in consultation with the Chairman of the Governing Board, the Chairman of the Board's Committee on Standards, Design and Methodology, staff of the National Center for Education Statistics, and the Executive Director of the Council of the Great City Schools.
7. The Executive Director of the Governing Board shall recommend new districts for participation in TUDA to the Governing Board for final action.
8. The Executive Director of the Governing Board shall send notification of the Board's decision regarding district participation in TUDA to the district and to the Commissioner of Education Statistics.

Potential Pool of Eligible Districts

The list of eligible districts shall be posted on the website of the National Assessment Governing Board (www.nagb.org) and made publicly available through other appropriate means. The list of districts will change from time to time due to changes in the population of the district and the district setting.