

National Assessment Governing Board

Executive Committee

Thursday, August 3, 2017

4:30 – 6:00 pm

AGENDA

4:30 – 4:35 pm Welcome and Agenda Overview
Terry Mazany, Chair

4:35 – 4:50 pm Governing Board Updates
Follow-up to Board Meeting Structure Options
Bill Bushaw, Executive Director

Policy Updates
Lily Clark, Assistant Director for Policy and Research

4:50- 5:00 pm **ACTION:** Nomination of Board Vice Chair for the Term October 1,
2017 – September 30, 2018
Terry Mazany
Ronnie Musgrove

5:00 – 5:30 pm **ACTION:** Strategic Vision #10 – Establish the Charge for the Ad Hoc Committee on Measures of Postsecondary Preparedness Attachment A
Terry Mazany

5:30 – 6:00 pm Long-Term Trend Discussion Attachment B
Joe Willhoft, Vice Chair, COSDAM

Resolution: The Executive Committee's Charge to the Committee on Measures of Postsecondary Preparedness

Whereas, on November 18, 2016, the National Assessment Governing Board unanimously approved the Strategic Vision to guide its work through the year 2020; and

Whereas, the Strategic Vision established a Board priority (SV#10) to “*Develop new approaches to measure the complex skills required for transition to postsecondary education and career*”; and

Whereas, on August 3, 2017, the Governing Board Chair created the ad hoc Committee on Measures of Postsecondary Preparedness to pursue this priority; and

Whereas, the Governing Board Chair tasked the Executive Committee to establish the charge to guide the ad hoc Committee on Measures of Postsecondary Preparedness;

Therefore, the Executive Committee resolves that:

- 1. The Committee on Measures of Postsecondary Preparedness shall review existing research, collect expert testimony, and prepare recommendations for the Governing Board's consideration to achieve Strategic Vision priority #10.**
- 2. While the current legislation guiding the National Assessment of Educational Progress (P.L. 107-279) should provide parameters for the approaches to accomplish this priority, the Committee on Measures of Postsecondary Preparedness may consider options that could require amendments to current legislation.**
- 3. The Committee on Measures of Postsecondary Preparedness will report its recommendations to the Governing Board no later than the November 2018 Board meeting.**

Long-Term Trend Discussion: Next Steps

At the August Executive Committee meeting, COSDAM Vice Chair Joe Willhoft will lead a discussion about the potential next steps for the Long-Term Trend (LTT) assessment. In this, the Board will consider the pros and cons of each of the following options regarding the future of LTT:

1. Transadapt Long-Term Trend from paper and pencil to digital-based assessments, produce assessment frameworks, perform a bridge study for each age group, and keep the assessments in the existing Long-Term Trend administration windows that do not align with main NAEP;
2. Ask Congress to remove the legislative requirement and cease administration of Long-Term Trend; or
3. Ask Congress to remove the legislative requirement to conduct the Long-Term Trend but perform a special study in which Long-Term Trend is administered one last time in an attempt to connect future main NAEP results with the long-standing Long-Term Trend lines.

No decision regarding LTT will be made by the Board at the August 2017 meeting. This Executive Committee discussion will inform the next steps to help the Board achieve priority #7 of the Strategic Vision to: “Research policy and technical implications related to the future of NAEP Long-Term Trend assessments in reading and mathematics.”

The following pages are provided as a refresher of the background on the LTT and a summary of the substantive work already conducted to inform the Board’s work on this priority, including commissioned papers to inform a symposium and subsequent presentation at the American Educational Research Association (AERA) on the topic of LTT.

Background

NAEP includes two national assessment programs—Long-Term Trend (LTT) NAEP and main NAEP. While both assessments enable NAEP to measure student progress over time, there are key differences between the two assessments. The NAEP LTT assessment measures national reading and mathematics performance at ages 9, 13 and 17. In contrast, the main NAEP assessments focus on populations of students defined by grade, rather than age, and go beyond the national level to provide results at the state level and for 27 urban districts. LTT trend lines date back to the early 1970s, and main NAEP trend lines start in the early 1990s. The content differs as well—for example, LTT mathematics measures more basic skills than the current main NAEP. Sample items can be found at: <https://nces.ed.gov/NationsReportCard/nqt/Search>.

The main NAEP assessments in reading and mathematics are administered every two years, as required by law. The administration of LTT assessments in reading and mathematics at ages 9, 13, and 17 is also required by law, but the periodicity is not specified. The NAEP LTT

assessment had been administered approximately every four years over the past two decades (and more frequently prior to that), but were last administered in 2012 and will be next administered in 2024. The Governing Board postponed the NAEP LTT planned administration for 2016 and 2020 due to budgetary constraints. Some stakeholders have expressed concern with the gap of 12 years between LTT administrations, which represents a cohort's entire length of schooling. On the other hand, there are stakeholders who argue that the NAEP LTT is less useful now that main NAEP provides trend information back to the early 1990s and that LTT should be eliminated.

In 2012, [the Future of NAEP panel](#) recommended exploring ways of consolidating or combining LTT and main NAEP data collections. This is a complex challenge due to the many differences in content, sampling, and administration of the assessments. To explore the feasibility of combining the data collection efforts and to debate the relative merits of NAEP LTT, the Governing Board organized a symposium on the future of NAEP Long-Term Trend. The symposium took place on March 2, 2017, immediately preceding the quarterly Governing Board meeting.

In advance of the symposium, Edward Haertel of Stanford University (who previously served as Chair of the Future of NAEP panel and, as a previous member of the Governing Board, chaired COSDAM) prepared a comprehensive white paper on the history of NAEP Long-Term Trend and a consideration of current issues. The paper was distributed to four additional participants, who each prepared a shorter response paper on their perspective of the future of NAEP LTT (*these papers are hyperlinked under each panel presentation below*). The papers were disseminated in advance of the symposium and served as the basis for discussion during the March 2nd event. In addition, the participants discussed their perspectives and solicited external input at a session during the annual American Educational Research Association (AERA) conference on April 29th. At both events, Acting NCES Commissioner Peggy Carr also participated and provided her perspective on the operational feasibility of the various options for the future of NAEP LTT.

Summary of the Symposium on the Future of NAEP Long-Term Trend

Introduction

Moderator Dr. Joe Willhoft thanked attendees for joining the discussion on the future of the National Assessment of Educational Progress (NAEP) Long-Term Trend assessment. He summarized the role of the National Assessment Governing Board and its relationship to the National Center for Education Statistics (NCES), which carries out the operations of NAEP.

Dr. Willhoft introduced and welcomed the symposium panel:

- Dr. Edward Haertel is a former member of the Governing Board and has served as president and chair of the National Research Council's Board of Testing and fellow of the American Educational Research Association and the American Psychological Association. He serves as an elected member of the National Academy of Education.
- Mr. Jack Jennings is a former president and CEO of the Center on Education Policy. He served as subcommittee staff director and as general counsel for the House of

Representatives Committee on Education and Labor. He is an elected member of the National Academy of Education.

- Dr. Lou Fabrizio is the director of data research and federal policy at the North Carolina Department of Public Instruction, serving as the state's federal liaison with the Education Department. He is a former member of the Governing Board and currently serves as a member of the NCES Advisory Task Force.
- Dr. Ina Mullis is a professor of educational research, measurement, and evaluation at Boston College and the executive director of the Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS) International Study Center. Previously, she was the project director of NAEP at Educational Testing Service (ETS). She currently serves on the NAEP Validity Studies Panel.
- Dr. Andrew Kolstad is a former senior technical advisor and psychometrician in the Assessment Division of NCES. He took the lead role in designing the 2004 bridge study for the NAEP Long-Term assessment. Through his consulting firm, P20 Strategies, he works with the Governing Board and for NAEP contractors.
- Dr. Peggy Carr is the Acting Commissioner of NCES. She oversees the collection, analysis, and reporting of education data from preschool through post-secondary education. She is one of the nation's foremost experts on student assessment.

Panel Presentations

Edward Haertel

[Future of NAEP Long-Term Trend Assessments](#)

Dr. Haertel presented the origins of the LTT assessment and how it differs from main NAEP. His slides included a graphic showing the decades-long span of the LTT assessment before main NAEP trends began. He reminded the audience that NAEP began as a series of exercises with no scale scores.

He explained that in 1983, ETS introduced item response theory, and scales were retrofitted to the old exercise pools. Frameworks continued to improve after the founding of the National Assessment Governing Board in 1988. A 2004 bridge study helped align the Long-Term Trend assessment with legislation and closer to the ideal of assessing all students, with the inclusion of more students with disabilities. Today, main NAEP is transitioning to a digital platform while the Long-Term Trend assessment remains a paper-and-pencil assessment.

Dr. Haertel argued for maintaining the LTT assessment and examining major actions required to ensure its viability in the future. He stated that main NAEP frameworks are updated periodically to reflect changing concepts and desired learning outcomes. The LTT assessment remains stable, making it a strong complement to main NAEP. It can also aid in policymaking because it is age-based, not grade-based, so the discrepancies can be analyzed between the two assessments. Tracking age-based cohorts can inform important policy questions on school enrollment (e.g., grade retention and grade acceleration).

Dr. Haertel encouraged greater use of the LTT assessment in the future. He stated that this will require modernization of the test specifications and frameworks, and it needs to be clear what the LTT is measuring alongside what main NAEP provides. Dr. Haertel emphasized the need for further bridge studies at each age level and a commitment to ongoing maintenance. He is confident that new uses for the assessment can be found through more study, increasing the value of the LTT assessment.

Jack Jennings

[Why Continue An Old Assessment?](#)

Mr. Jennings thanked the Governing Board for being thoughtful in the process and discussion of the LTT assessment.

Mr. Jennings outlined his reasons for continuing and strengthening the LTT assessment. First, he stated, it is the law. Second, he believes delaying the assessment for 12 years jeopardizes the usefulness of the data. He suggested that the LTT assessment could be made more useful by more frequent administration.

Mr. Jennings stated that the LTT assessment serves as a safeguard for the deficiencies of main NAEP, including the controversies associated with achievement levels. He stated that his background is not in testing; it is policy, so he supports long-term assessment as a valuable tool in policy and communication.

Mr. Jennings stated that he understood the ongoing concern of cost, which was presented in many of the response papers. He is in favor of searching for ways to retain the essence of the LTT assessment while being mindful of the cost of administration.

Lou Fabrizio

[Is It Time to Retire Long-Term Trend?](#)

Dr. Fabrizio explained his role in the state of North Carolina and his experience with policy, legislation, and the testing and accountability program. He approached the issue from a state perspective. He acknowledged that the Long-Term Trend assessment is currently the law, but stated that the U.S. Congress should remove the requirement to administer this assessment.

Dr. Fabrizio stated that the Governing Board should pursue all efforts to enhance main NAEP, including increasing participation by private schools.

Dr. Fabrizio said the LTT assessment is out of date and does not provide actionable information at the state level. He is also in favor of reducing redundant testing. He feels the resources used for Long-Term Trend assessment could be better spent with main NAEP assessments that offer states help with benchmarking.

Ina Mullis

[Content of the Long-Term Trend Assessments Compared to Main NAEP](#)

Dr. Mullis began her presentation with a comparison of the Long-Term Trend assessment and main NAEP. She then explained her role in the international assessment community.

Dr. Mullis stated that main NAEP is a world-class assessment with challenging content and forward-looking frameworks. She felt that the Long-Term Trend assessment, on the other hand, is outdated in both content and measurement methods. LTT is almost wholly comprised of multiple choice items, the LTT reading passages are short and inauthentic, and LTT mathematics includes little or no problem solving questions. Dr. Mullis believes that main NAEP includes more challenging material than the Long-Term Trend assessment, which focuses on basic skills. There is overlap among items across ages 9, 13, and 17, which puts a ceiling on item difficulty of the LTT assessment.

She also stated that there are no formal frameworks for the LTT, and it is hard to justify using scarce resources to support the assessment. The high-quality main NAEP has nearly 25 years of trends now; therefore, she believes the Long-Term Trend assessment is no longer needed.

Andrew Kolstad

[A Rescue Plan for the NAEP Long-Term Trend Assessments: Thoughts on Edward Haertel's White Paper](#)

Mr. Kolstad explained that his paper focused on the technical point of view. He did not advocate in favor or against the Long-Term Trend assessment. He stated that he was concerned with the 2024 assessment as it is currently planned. He titled his paper “A Rescue Plan” because it outlines issues that must be addressed for the assessment’s future success, including the possibility of integrating the LTT and main NAEP or switching to a digitally-based format.

Mr. Kolstad addressed the issue of creating a common testing window between main NAEP and the LTT assessment to preserve resources. For instance, he noted that a separate scoring operation would no longer be needed and managerial costs could be conserved. He proposed that a bridge study could examine the feasibility of this merger.

He echoed Ina Mullis’ caution about the importance of understanding what the LTT assessment is really measuring.

Peggy Carr

Dr. Carr began by discussing the mode issue—there is both paper-and-pencil and digital-based assessments; the window issue—the Long-Term Trend assessment occurs during a different window of time during the year than main NAEP; the length issue—the item blocks differ between LTT and main NAEP; and the test content issue, which she thinks was addressed well in the panelists’ papers.

She pointed out the importance of moving LTT into the digital space because support for paper-and-pencil administration is disappearing. She stated that the costs associated with the bridge

studies are significant and should not wait until the 2024 timeframe. From her perspective as the operational leader of all the NAEP programs, she believes “transadapting” the Long-Term Trend assessment is the only option. Dr. Carr explained that transadapting means taking a paper-and-pencil assessment and adapting it to the digital environment to take advantage of the features that one can use to enhance it without changing the measurement of the underlying construct the item intends to assess.

Dr. Carr stated that NCES should examine the efficiency of changing the length of the LTT assessment to allow for the possibility of administering it with main NAEP. She presented options on how to adapt LTT and main NAEP assessments for cost and scoring challenges. She added that she was concerned about the platform used to administer the assessments and how they can be used in the same session. For instance, the Long-Term Trend assessment consists of three 15-minute blocks, while main NAEP in the digital world is now 30 minutes per block. This is an important consideration for some of the designs that have been recommended. Dr. Carr also expressed uncertainty about what platform would be used to administer the assessments. She stated that the bridge study would need to address multiple factors, including the mode, the windows, and the test length.

Dr. Carr described the communications challenges of explaining to stakeholders what the Long-Term Trend assessment is measuring and the confusion by the public in interpreting the results when both LTT and main NAEP are released in close proximity to each other.

She also stated that the budget is an ongoing concern in whatever plan is chosen. Moving the Long-Term Trend assessment to the main NAEP window has costs associated and should be addressed when considering different options. Moving the Long-Term Trend assessment into main NAEP would be expensive. Currently, there are many assumptions about spiraling them together, and these need to be studied. She stated that she does not have a preference and is examining administrative and operational challenges with all options.

In subsequent discussions, Dr. Carr noted that the LTT Reading assessment faces an additional challenge. The last time this assessment was administered in 2012, there were challenges in scoring, particularly for the highest level responses. Students were giving long, elaborate responses that were not in the scoring rubric, which presented difficulty to the scorers. Because of such issues, it might be that differences in how students are learning and responding to questions could threaten the integrity of maintaining trend lines, at least for reading.

Discussion

The following comments and questions were raised by audience members:

- As a policymaker, the role of the Long-Term Trend assessment is minimal and main NAEP is extensive. The ongoing value of the Long-Term Trend assessment is questionable. There was agreement with Dr. Mullis and her statement that digital assessment and content are imperative. The Long-Term Trend assessment must add something beyond main NAEP to provide value in policy decisions.

- What is the relevance of NAEP to a parent who wants to know whether his or her school is doing a good job? Do NAEP scores reflect what the school is doing right or wrong?
- If NAEP is used for policy decisions, what is being done to engage other stakeholders including industry, parents, students, and educators in the policy decisions? How are others part of the decision to continue the Long-Term Trend assessment and understand what it is measuring?
- There is an issue of cost/benefit analysis. Given the unexpected dips in the 2015 main NAEP results, it would have been helpful to have a second set of reading and math results in 2016 from LTT. And what are the costs associated with other NAEP investments, including the Technology and Engineering Literacy assessment and the Trial Urban District Assessment program in several dozen large cities?
- A local superintendent of a major urban district pays close attention to main NAEP but mostly ignores the Long-Term Trend assessment. Given the demographic shift of students taking the assessments today versus the demographics of 45 years ago, could those shifts change the relevance of the Long-Term Trend assessment?
- Main NAEP has useful information that schools and parents can use on some levels, including data about in- and out-of-school factors that affect learning, while the Long-Term Trend assessment lacks this context.
- What is the Governing Board doing to address what cognitive science says about how students solve math problems and how that compares with main NAEP and the LTT?

AERA Session on the Future of NAEP Long-Term Trend

On April 29, 2017, the same panel presentations were delivered at an invited session of the American Educational Research Association (AERA) annual meeting in San Antonio, Texas. Dr. Ina Mullis was unable to attend; her paper was presented by Dr. Joe Willhoft.

The following comments and questions were raised by attendees:

- What is the decision process for the Governing Board?
- Is it important if we lose current knowledge of students' basic skills in mathematics and reading as measured in the Long-Term Trend assessment?
- Congress is not likely to argue with cutting something that would reduce a budget.
- What about the validity of policy inferences or usefulness of decisions made with the NAEP assessment program?