

Response to Evaluation of NAEP Achievement Levels for Mathematics and Reading

Background

Public Law 107-279 states:

The achievement levels shall be used on a trial basis until the Commissioner for Education Statistics determines, as a result of an evaluation under subsection (f), that such levels are reasonable, valid, and informative to the public.

Even after being in use for about 25 years and undergoing previous evaluations (1993, 1998, 2009), the NAEP achievement levels are still considered to be on a trial basis. During his tenure as NCES Commissioner, Jack Buckley initiated a new evaluation to determine whether the trial status could be resolved.

About the Evaluation

The National Center for Education Evaluation and Regional Assistance (NCEE), part of the Institute for Education Sciences (IES), administered the recent evaluation of the NAEP achievement levels. On September 29, 2014, NCEE awarded a contract to The National Academy of Sciences, Engineering, and Medicine to perform this work.

Objectives for the evaluation include the following:

- Determine how "reasonable, valid, reliable and informative to the public" will be operationalized in this study.
- Identify the kinds of objective data and research findings that will be examined.
- Review and analyze extant information related to the study's purpose.
- Gather other objective information from relevant experts and stakeholders, without creating burden for the public through new, large-scale data collection.
- Organize, summarize, and present the findings from the evaluation in a written report, including a summary that is accessible for nontechnical audiences, discussing the strengths/weaknesses and gaps in knowledge in relation to the evaluation criteria.
- Provide, prior to release of the study report, for an independent external review of that report for comprehensiveness, objectivity, and freedom from bias.
- Plan and conduct dissemination events to communicate the conclusions of the final report to different audiences of stakeholders.

Design

This study focused on the achievement levels used in reporting NAEP results for the reading and mathematics assessments in grades 4, 8, and 12. Specifically, the study reviewed developments over the past decade in the ways achievement levels for NAEP were set and used and evaluated whether the resulting achievement levels are "reasonable, valid, reliable, and informative to the public." The study relied on an independent committee of experts with a broad range of expertise related to assessment, statistics, social science, and education policy. The project received oversight from the Board on Testing and Assessment (BOTA) and the Committee on National Statistics (CNSTAT) of the National Research Council.

Members of the interdisciplinary review committee were:

Name	Affiliation
Dr. Christopher F. Edley, Jr. (Chair)	University of California at Berkeley
Dr. Peter Afflerbach	University of Maryland, College Park
Dr. Sybilla Beckmann	University of Georgia
Dr. H. Russell Bernard	University of Florida
Dr. Karla Egan	EdMetric LLC
Dr. David J. Francis	University of Houston
Dr. Margaret E. Goertz	University of Pennsylvania
Dr. Laura Hamilton	The RAND Corporation
Dr. Brian W. Junker	Carnegie Mellon University
Dr. Suzanne Lane	University of Pittsburgh
Ms. Sharon J. Lewis	Retired (formerly with the Council of the Great City Schools)
Dr. Bernard L. Madison	University of Arkansas
Dr. Scott Norton	Council of Chief State School Officers
Dr. Sharon Vaughn	The University of Texas at Austin
Dr. Lauress L. Wise	HumRRO

The final report was released on November 17, 2016; a free PDF of the full report can be downloaded at: <https://www.nap.edu/catalog/23409/evaluation-of-the-achievement-levels-for-mathematics-and-reading-on-the-national-assessment-of-educational-progress>. The Governing Board received a briefing from staff at the National Academy of Sciences, Engineering, and Medicine and members of the interdisciplinary review committee during the most recent quarterly Board meeting on November 19, 2016.

Governing Board Response and Next Steps

As stated in the NAEP legislation, the Commissioner of NCES is to use the findings from the evaluation to decide whether the achievement levels should continue to be used on a "trial basis" or whether that designation can be removed. In addition, the final report included conclusions

and recommendations that have implications for future Governing Board achievement levels-setting activities. Public Law 107-279 specifies that the Governing Board must prepare a formal response to the evaluation:

Not later than 90 days after an evaluation of the student achievement levels under section 303(e), the Assessment Board shall make a report to the Secretary, the Committee on Education and the Workforce of the House of Representatives, and the Committee on Health, Education, Labor, and Pensions of the Senate describing the steps the Assessment Board is taking to respond to each of the recommendations contained in such evaluation.

Due to the timing of the evaluation report release, the 90 day window concluded prior to the March 2017 Governing Board meeting. Therefore, on November 19, 2016, the Board granted a joint delegation of authority to COSDAM and the Executive Committee for formal approval of the report to the Secretary, the Committee on Education and the Workforce of the House of Representatives, and the Committee on Health, Education, Labor, and Pensions of the Senate describing the steps the Governing Board is taking to respond to each of the recommendations contained in the evaluation.

COSDAM met via teleconference on December 9, 2016 to discuss an initial draft response to the evaluation. On December 19, 2016, the Executive Committee and COSDAM met to discuss and take action on a revised response. The final response was approved by a vote of 9-0 with one abstention. The response was sent to Secretary John King, the Committee on Education and the Workforce of the House of Representatives, and the Committee on Health, Education, Labor, and Pensions of the Senate on December 20, 2016.

The Governing Board response refers to several new activities to be undertaken, in addition to plans to update the current policy on [Developing Student Performance Levels for NAEP](#). Much of the work aligns with the Strategic Vision and can be performed in collaboration with NCES. Ongoing discussions with COSDAM and the full Board will take place over the next several quarterly meetings to plan and implement the recommendations from the evaluation.

National Assessment Governing Board's Response to the National Academies of Sciences, Engineering, and Medicine 2016 Evaluation of NAEP Achievement Levels

Legislative Authority

Pursuant to the National Assessment of Educational Progress (NAEP) legislation (Public Law 107-279), the National Assessment Governing Board (hereafter the Governing Board) is pleased to have this opportunity to apprise the Secretary of Education and the Congress of the Governing Board response to the recommendations of the National Academies of Sciences, Engineering, and Medicine evaluation of the NAEP achievement levels for mathematics and reading (Edley & Koenig, 2016).

The cited legislation charges the Governing Board with the authority and responsibility to “develop appropriate student achievement levels for each grade or age in each subject area to be tested.” The legislation also states that “such levels shall be determined by... a national consensus approach; used on a trial basis until the Commissioner for Education Statistics determines, as a result of an evaluation under subsection (f), that such levels are reasonable, valid, and informative to the public; ... [and] shall be updated as appropriate by the National Assessment Governing Board in consultation with the Commissioner for Education Statistics” (Public Law 107-279).

Background

NAEP is the largest nationally representative and continuing assessment of what our nation's elementary and secondary students know and can do. Since 1969, NAEP has been the country's foremost resource for measuring student progress and identifying differences in student achievement across student subgroups. In a time of changing state standards and assessments, NAEP serves as a trusted resource for parents, teachers, principals, policymakers, and researchers to compare student achievement across states and select large urban districts. NAEP results allow the nation to understand where more work must be done to improve learning among all students.

For 25 years, the NAEP achievement levels (*Basic*, *Proficient*, and *Advanced*) have been a signature feature of NAEP results. While scale scores provide information about student achievement over time and across student groups, achievement levels reflect the extent to which student performance is “good enough,” in each subject and grade, relative to aspirational goals.

Since the Governing Board began setting standards in the early 1990s, achievement levels have become a standard part of score reporting for many other assessment programs in the US and abroad.

Governing Board Response

Overview

The Governing Board appreciates the thorough, deliberative process undertaken over the past two years by the National Academies of Science, Engineering, and Medicine and the expert members of the Committee on the Evaluation of NAEP Achievement Levels for Mathematics and Reading. The Governing Board is pleased that the report concludes that the achievement levels are a meaningful and important part of NAEP reporting. The report states that, “during their 24 years [the achievement levels] have acquired meaning for NAEP’s various audiences and stakeholders; they serve as stable benchmarks for monitoring achievement trends, and they are widely used to inform public discourse and policy decisions. Users regard them as a regular, permanent feature of the NAEP reports” (Edley & Koenig, 2016; page Sum-8). The Governing Board has reviewed the seven recommendations presented in the report and finds them reasonable and thoughtful. The report will inform the Board’s future efforts to set achievement levels and communicate the meaning of NAEP *Basic*, *Proficient*, and *Advanced*. The recommendations intersect with two Governing Board documents, the Strategic Vision and the achievement levels policy, described here.

On November 18, 2016, the Governing Board adopted a Strategic Vision (<https://www.nagb.org/content/nagb/assets/documents/newsroom/press-releases/2016/nagb-strategic-vision.pdf>) to guide the work of the Board through 2020, with an emphasis on innovating to enhance NAEP’s form and content and expanding NAEP’s dissemination and use. The Strategic Vision answers the question, “How can NAEP provide information about how our students are doing in the most innovative, informative, and impactful ways?” The Governing Board is pleased that several of the report recommendations are consistent with the Board’s own vision. The Governing Board is committed to measuring the progress of our nation’s students toward their acquisition of academic knowledge, skills, and abilities relevant to this contemporary era.

The Governing Board’s approach to setting achievement levels is articulated in a policy statement, “Developing Student Performance Levels for the National Assessment of Educational Progress” (<https://www.nagb.org/content/nagb/assets/documents/policies/developing-student-performance.pdf>). The policy was first adopted in 1990 and was subsequently revised in 1995, with minor wording changes made in 2007. The report motivates the revision of this policy, to add clarity and intentionality to the setting and communication of NAEP achievement levels.

The seven recommendations and the Governing Board response comprise a significant research and outreach trajectory that the Governing Board can pursue over several years in conjunction with key partners. The Governing Board will implement these responses within resource constraints and in conjunction with the priorities of the Strategic Vision.

Evaluating the Alignment of NAEP Achievement Level Descriptors

Recommendation #1: Alignment among the frameworks, the item pools, the achievement-level descriptors, and the cut scores is fundamental to the validity of inferences about student achievement. In 2009, alignment was evaluated for all grades in reading and for grade 12 in mathematics, and changes were made to the achievement-level descriptors, as needed. Similar research is needed to evaluate alignment for the grade 4 and grade 8 mathematics assessments and to revise them as needed to ensure that they represent the knowledge and skills of students at each achievement level. Moreover, additional work to verify alignment for grade 4 reading and grade 12 mathematics is needed.

The report's primary recommendation is to evaluate the alignment, and revise if needed, the achievement level descriptors for NAEP mathematics and reading assessments in grades 4, 8, and 12. The Governing Board intends to issue a procurement for conducting studies to achieve this goal. The Governing Board has periodically conducted studies to evaluate whether the achievement level descriptors in a given subject should be revised, based on their alignment with the NAEP framework, item pool, and cut scores. The Governing Board agrees that this is a good time to ensure that current NAEP mathematics and reading achievement level descriptors align with the knowledge and skills of students in each achievement level category. In conjunction with the response to Recommendation #3, the updated Board policy on NAEP achievement levels will address the larger issue of specifying a process and timeline for conducting regular recurring reviews of the achievement level descriptions in all subjects and grades.

The Governing Board agrees strongly with the recommendation that, while evaluating alignment of achievement level descriptors is timely, it is not necessary to consider changing the cut scores or beginning a new trend line at this time. The NAEP assessments are transitioning from paper-based to digital assessments in 2017, and current efforts are focused on ensuring comparability between 2015 and 2017 scores. The Governing Board articulated this in the 2015 Resolution on Maintaining NAEP Trends with the Transition to Digital-Based Assessments (<https://www.nagb.org/content/nagb/assets/documents/policies/resolution-on-trend-and-dba.pdf>).

Recommendation #2: Once satisfactory alignment among the frameworks, the item pools, the achievement-level descriptors, and the cut scores in NAEP mathematics and reading has been demonstrated, their designation as trial should be discontinued. This work should be completed

and the results evaluated as stipulated by law: (20 U.S. Code 9622: National Assessment of Educational Progress: <https://www.law.cornell.edu/uscode/text/20/9622> [September 2016]).

Ultimately, the Commissioner of Education Statistics is responsible for determining whether the “trial” designation is removed. The Governing Board is committed to providing the Commissioner with the information needed to make this determination in an expedient manner.

Regular Recurring Reviews of the Achievement Level Descriptors

Recommendation #3: To maintain the validity and usefulness of achievement levels, there should be regular recurring reviews of the achievement-level descriptors, with updates as needed, to ensure they reflect both the frameworks and the incorporation of those frameworks in NAEP assessments.

The Board’s current policy on NAEP achievement levels contains several principles and guidelines for *setting* achievement levels but does not address issues related to the continued use or reporting of achievement levels many years after they were established. The revised policy will seek to address this gap by including a statement of periodicity for conducting regular recurring reviews of the achievement level descriptors, with updates as needed, as called for in this recommendation. The Governing Board agrees that it is important to articulate a process and timeline for conducting regular reviews of the achievement level descriptors rather than performing such reviews on an ad hoc basis.

Relationships Between NAEP Achievement Levels and External Measures

Recommendation #4: Research is needed on the relationships between the NAEP achievement levels and concurrent or future performance on measures external to NAEP. Like the research that led to setting scale scores that represent academic preparedness for college, new research should focus on other measures of future performance, such as being on track for a college-ready high school diploma for 8th-grade students and readiness for middle school for 4th-grade students.

In addition to the extensive work that the Governing Board has conducted at grade 12 to relate NAEP mathematics and reading results to academic preparedness for college, the Governing Board has begun research at grade 8 with statistical linking studies of NAEP mathematics and reading and the ACT Explore assessments in those subjects. This work was published while the evaluation was in process and was not included in the Committee’s deliberations. Additional studies in NAEP mathematics and reading at grades 4 and 8 are beginning under contract to the National Center for Education Statistics (NCES). The Governing Board’s Strategic Vision includes an explicit goal to increase opportunities for connecting NAEP to other national and international assessments and data. Just as the Board’s previous research related grade 12 NAEP

results in mathematics and reading to students' academic preparedness for college, the Governing Board anticipates that additional linkages with external measures will help connect the NAEP achievement levels and scale scores to other meaningful real-world indicators of current and future performance.

Interpretations and Uses of NAEP Achievement Levels

Recommendation #5: Research is needed to articulate the intended interpretations and uses of the achievement levels and collect validity evidence to support these interpretations and uses. In addition, research to identify the actual interpretations and uses commonly made by NAEP's various audiences and evaluate the validity of each of them. This information should be communicated to users with clear guidance on substantiated and unsubstantiated interpretations.

The Governing Board's Strategic Vision emphasizes improving the use and dissemination of NAEP results, and the Board's work in this area will include achievement levels. The Governing Board recognizes that clarity and meaning of NAEP achievement levels (and scale scores) are of utmost importance. The Governing Board will issue a procurement to conduct research to better understand how various audiences have used and interpreted NAEP results (including achievement levels). The Governing Board will work collaboratively with NCES to provide further guidance and outreach about appropriate and inappropriate uses of NAEP achievement levels.

Guidance for Inferences Made with Achievement Levels versus Scale Scores

Recommendation #6: Guidance is needed to help users determine inferences that are best made with achievement levels and those best made with scale score statistics. Such guidance should be incorporated in every report that includes achievement levels.

The Governing Board understands that improper uses of achievement level statistics are widespread in the public domain and extend far beyond the use of NAEP data. Reports by the Governing Board and NCES have modeled appropriate use of NAEP data and will continue to do so. This recommendation is also consistent with the goal of the Strategic Vision to improve the dissemination and use of NAEP results. The Governing Board will continue to work with NCES and follow current research to provide guidance about inferences that are best made with achievement levels and those best made with scale score statistics.

Regular Cycle for Considering Desirability of Conducting a New Standard Setting

Recommendation #7: NAEP should implement a regular cycle for considering the desirability of conducting a new standard setting. Factors to consider include, but are not limited to: substantive changes in the constructs, item types, or frameworks; innovations in the modality for administering assessments; advances in standard setting methodologies; and changes in the policy environment for using NAEP results. These factors should be weighed against the downsides of interrupting the trend data and information.

When the Board's achievement levels policy was first created and revised in the 1990s, the Board was setting standards in each subject and grade for the first time and had not yet considered the need or timeline for re-setting standards. To address this recommendation, the Governing Board will update the policy to be more explicit about conditions that require a new standard setting.

Board's Commitment

The Governing Board remains committed to its congressional mandate to set "appropriate student achievement levels" for the National Assessment of Educational Progress. The Board appreciates the report's affirmation that NAEP achievement levels have been set thoughtfully and carefully, consistent with professional guidelines for standard setting, and based on extensive technical advice from respected psychometricians and measurement specialists. The Board also takes seriously the charge to develop the current achievement levels through a national consensus approach, involving large numbers of knowledgeable teachers, curriculum specialists, business leaders, and members of the general public throughout the process. This is only fitting given the Governing Board's own congressionally mandated membership that explicitly includes representatives from these stakeholder groups.

The Governing Board remains committed to improving the process of setting and communicating achievement levels. The Governing Board is grateful for the report recommendations that will advance these aims.

Reference

Edley, C. & Koenig, J. A. (Ed.). (2016). *Evaluation of the Achievement Levels for Mathematics and Reading on the National Assessment of Educational Progress*. Washington, DC: The National Academies Press.