

## **Discussion of Revised Draft Policy on NAEP Achievement Level Setting (SV #5)**

For the past 1.5 years, COSDAM members have been discussing the need to revise the 1995 Governing Board policy on [Developing Student Performance Levels for NAEP](#). The Board’s formal response to the November 2016 evaluation of the NAEP achievement levels (attached) noted that several of the report recommendations would be addressed through a revision of the Board policy. In particular, the Board’s response stated that the updated policy will specify a process and timeline for conducting regularly recurring reviews of the achievement level descriptions (ALDs) and will be explicit about the conditions that necessitate consideration of a new standard setting. In addition, one of the planned activities for the implementation of the Strategic Vision is to consider new approaches to creating and updating the achievement level descriptions in the revision of the Board policy on achievement levels.

Given that the policy is over 20 years old, there is also a need to revisit the policy more generally to ensure that it reflects current best practices in standard setting.

Several activities have informed the draft revised policy, including ongoing COSDAM discussion, input from technical experts in standard setting, and several research efforts. The March and May 2018 COSDAM meetings were entirely devoted to this topic.

The table below summarizes both prior and planned upcoming activities:

<b>Activity</b>	
Initial COSDAM discussions of planned activities and timeline	March – May 2017
Initial input from standard setting experts	March – April 2017
Initial full Board discussion about potential elements of policy revision	August 2017
Technical advisory panel to seek expert advice and debate on best practices in achievement level setting	January 2018
Literature review of best practices for creating and updating ALDs	February 2018
Technical memo on developing a validity argument for the NAEP achievement levels	February 2018
COSDAM discussion of using research to inform goals for policy revision	March 2018
COSDAM review and discussion of draft revised policy	May 2018
Joint COSDAM/R&D discussion on communicating achievement levels	May 2018
COSDAM call to finalize a consensus draft for full Board discussion	June 2018
<b>Full Board review and discussion of draft revised policy</b>	<b>August 2018</b>
Public comment on draft revised policy	Fall 2018
Call to discuss any additional revisions to draft policy	Fall 2018
Planned full Board action on revised policy	November 2018

Compared to the current (1995) policy on Developing Student Performance Levels for NAEP, the attached proposed revised policy reflects:

- Reorganization of principles, streamlining of language, minimization of redundancies
- Minor (non-substantive) edits to the NAEP policy definitions for clarity
- A change in terminology from *Proficient* to *NAEP Proficient* to better differentiate the NAEP achievement levels from other common uses of *Basic, Proficient, Advanced*
- A new principle on periodic review of achievement level descriptions and cut scores, prompted by the Board's response to the evaluation of NAEP achievement levels
- A new principle to clarify participation of multiple stakeholders at various points throughout process
- A new principle to summarize the role of the Board
- Reference to an interpretative guide that would accompany the release of NAEP results and explain how the achievement levels should (and should not) be used
- Reference to multiple types of achievement level descriptions (ALDs), including reporting ALDs that would be created using empirical data and written in terms of what students *do know and can do* rather than what students *should know and be able to do*
- Clarification on the standard setting participants, in particular the non-educator group
- Additional details about the achievement level setting process, including some practices that have become institutionalized over time (e.g., the use of "impact data")
- Removal of details on implementation directed to staff and contractors, which will instead be included in a "procedures manual"

On Saturday morning, August 4, 2018, COSDAM Chair Andrew Ho will briefly describe the draft revised policy, followed by full Board discussion. No action is intended at this time. Following the full Board discussion, additional edits will be incorporated and public comment will be sought. It is anticipated that the Board will take action on the revised policy statement during the November 2018 Board meeting.



## **National Assessment Governing Board**

### **Developing Student Achievement Levels for the National Assessment of Educational Progress**

#### **Policy Statement**

It is the policy of the National Assessment Governing Board to conduct a comprehensive, inclusive, and deliberative process to develop and review student achievement levels for the National Assessment of Educational Progress (NAEP). Achievement levels consist of general policy definitions for the *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced* levels, specific achievement level descriptions (ALDs) for each assessment, cut scores that demarcate adjacent levels, and exemplar items or tasks that illustrate performance at each level. This process shall be conducted according to widely accepted professional standards, to produce results that are reasonable, useful, and informative to the public.

The Governing Board, through its Committee on Standards, Design and Methodology (COSDAM), shall monitor the development and review of student achievement levels to ensure that the final Governing Board-adopted achievement level descriptions, cut scores, and exemplars comply with all principles of this policy.

The achievement level setting process shall be carried out by contractors selected through a competitive bidding process. The process shall be managed in a technically sound, efficient, cost-effective manner, and shall be completed in a timely fashion.

#### **Introduction**

Since its creation by Congress in 1988, the Governing Board has been responsible for developing student achievement levels for NAEP assessments. The Governing Board has carried out this important statutory responsibility by engaging with a broad spectrum of stakeholders to develop student achievement levels.

Under provisions of the National Assessment of Educational Progress Authorization Act of 2002 (P.L. 107-279), Congress authorized the Governing Board to, develop, “achievement levels that are consistent with relevant widely accepted

professional assessment standards and based on the appropriate level of subject matter knowledge for grade levels to be assessed” (Section 303(e)(2)(A)(i)(II).

Given this mandate, the Governing Board must ensure that all achievement level setting processes align with current best practices in standard setting, and that appropriate validity evidence is collected and documented to support the intended uses and interpretations of NAEP achievement levels.

The Governing Board has established the following **policy definitions** for the NAEP achievement levels, as expectations of what students should know and be able to do. They shall be consistent across all assessments in which achievement levels are set.

***NAEP Basic***

*This level denotes partial mastery of prerequisite knowledge and skills that are fundamental for performance at the NAEP Proficient level.*

***NAEP Proficient***

*This level represents solid academic performance for each NAEP assessment. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real world situations, and analytical skills appropriate to the subject matter.*

***NAEP Advanced***

*This level signifies superior performance beyond NAEP Proficient.*

The Governing Board engages multiple stakeholders throughout the achievement level setting process, including:

- |                             |                                   |
|-----------------------------|-----------------------------------|
| Teachers                    | Policymakers                      |
| Curriculum Experts          | Business Representatives          |
| Content Experts             | Parents                           |
| Assessment Specialists      | Users of Assessment Data          |
| State Administrators        | Researchers and Technical Experts |
| Local School Administrators | Members of the Public             |

This policy also complies with the documents listed below which express widely accepted technical and professional standards for achievement level setting. These standards reflect the agreement of recognized experts in the field, as well as the policy positions of major professional and technical associations concerned with educational testing.

In conjunction with this policy the Board shall maintain a procedures manual to establish and document additional details about how this policy is to be implemented. As professional standards evolve and new consensus documents are released, this policy and the procedures manual shall be updated to the extent that new professional standards require. Resources for this purpose shall include, but not be limited to the following:

*The Standards for Educational and Psychological Testing.* (2014). Washington, DC: American Educational Research Association, American Psychological Association, and National Council on Measurement in Education;

*Code of Fair Testing Practices in Education.* (2004). Washington, DC: Joint Committee on Testing Practices;

*Educational Measurement (4<sup>th</sup> ed.).* (2006). R.L. Brennan (Ed.). Westport, CT: Praeger; and

*National Center for Education Statistics (NCES) Statistical Standards.* (2012).

## **Principles for Setting Achievement Levels**

**Principle 1: Elements of Achievement Levels**

**Principle 2: Development of Achievement Level Recommendations**

**Principle 3: Validation and Reporting of Achievement Level Results**

**Principle 4: Periodic Review of Achievement Levels**

**Principle 5: Stakeholder Input**

**Principle 6: Role of the Governing Board**

## **Principle 1: Elements of Achievement Levels**

**The Governing Board is responsible for developing student achievement levels for each NAEP assessment. Achievement levels for each NAEP assessment consist of content achievement level descriptions (ALDs), cut scores that demarcate adjacent levels, and exemplar items or tasks that illustrate performance at each level.**

- a) Content achievement level descriptions (ALDs) translate the policy definitions into specific expectations about student knowledge and skills in a particular content area, at each achievement level, for each subject and grade. Content ALDs provide descriptions of specific expected knowledge, skills, or abilities of students performing at each achievement level. Content ALDs reflect the range of performance that items and tasks should measure. During the achievement level setting process, the purpose of content ALDs is to provide consistency and specificity for panelist interpretations of policy definitions for a given assessment. During reporting, content ALDs communicate the specific knowledge and skills represented by *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced* for a given assessment.
- b) Cut scores mark the minimum threshold score, the lower bound, for each achievement level. Performance within a given achievement level begins at the cut score for that level and ends just below the cut score for the successive achievement level.
- c) Exemplar items or tasks, including student responses, illustrate student performance within each of the achievement levels. They provide specific examples to help the public better understand what students in each achievement level can do.

## **Principle 2: Development of Achievement Level Recommendations**

**The Governing Board shall develop student achievement levels for NAEP, consistent with relevant widely accepted professional assessment standards, based on the appropriate level of subject matter knowledge.**

- a) A Design Document shall be developed at the beginning of the achievement level setting process, to describe in detail the scope of the achievement level setting project being undertaken, including but not limited to all planned materials, procedures, and analyses needed for the project. The Design Document shall be posted for public review with sufficient time to allow for a response from those who wish to provide one.
- b) The development of content achievement level descriptions (ALDs) shall be completed initially through the process that develops the assessment frameworks. (See the Governing Board Policy on Framework Development for additional details). The Board may then review and revise content ALDs to advance the purposes they serve, whether that is guiding an achievement level setting or informing the public about the meaning of achievement levels. Whether revised or not, the ALDs that guide

achievement level setting shall be articulated in terms of what students *should know and be able to do*. There shall be no content ALDs developed for performance below the *NAEP Basic* level.

- c) An achievement-level setting panel of subject matter experts shall be convened to recommend achievement level cut scores and exemplars.
- i. Each panel shall reflect diversity in terms of gender, race/ethnicity, region of the country, urbanicity, and experience with students with disabilities and English language learners. To ensure that they are qualified to make the judgments required by the achievement level setting process, individual panel members shall have expertise and experience in the specific content area in which the levels are being developed, expertise and experience in the education of students at the grade under consideration, and a general knowledge of assessment, curriculum, and student performance.
  - ii. This panel shall include teachers, non-teacher educators, and other interested members of the general public with relevant educational background and experience. Teachers shall comprise the majority of the panel, with non-teacher educators (e.g., curriculum directors, academic coaches, principals) accounting for no more than half the number of teachers. The remaining panelists shall be non-educators who represent the perspectives of additional stakeholders representing the general public, including parents, researchers, and employers.
  - iii. The size of the panels shall reflect best practice in standard setting and be operationally feasible while being large enough to allow for split panels. Most NAEP achievement level settings have historically included approximately 20-30 panelists per grade, divided into two comparable groups with a subset of shared items.
- d) Panelists shall receive training on all aspects of the achievement levels setting process to ensure that panelists are well-prepared to perform the achievement level setting tasks required of them. Panelists shall be instructed that their role is to make achievement level recommendations to the Governing Board. Training shall include but not be limited to: the purpose and significance of setting achievement levels for NAEP; the NAEP assessment framework for the given subject area; and administration of a sample assessment under NAEP-like conditions that students experience. It is important for panelists to arrive at a common conceptualization of *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced* based on the content ALDs. Panelists shall be trained on each element of the judgmental task they perform, including the selection of exemplar items. They should be led by capable *content facilitators* (who are content experts and have previous experience with achievement level setting) and *process facilitators* (who have background in standard setting and experience leading panelists through the achievement level setting process). Facilitators shall take a neutral stance and not attempt to influence panelist judgments.



- e) The achievement level setting method that generates cut score recommendations may differ depending upon the specific assessment. Nevertheless, the method shall have a solid research base and be appropriate for the content area, item types, number of items, scoring rubrics, and mode of administration, as applicable.
- f) Evaluations shall be administered to panelists throughout the achievement level setting process, in accordance with current best practices. Evaluations shall be part of every major component of the process, and panelists shall be asked to confirm their readiness for performing their tasks. Evaluation data may be used for formative purposes (to improve training and procedures in future meetings); summative purposes (to evaluate how well the process was conducted and provide procedural validity evidence); and to inform the Governing Board of any relevant information that could be useful when considering cut score recommendations. The panelists shall have an opportunity to indicate to the Board whether they believe the recommended cut scores are reasonable.
- g) In accordance with current best practices, feedback shall be provided to panelists, including “impact data” (i.e., the implications of their selected cut scores on the reported percentages of students at or above each achievement level).
- h) The process shall consist of at least two achievement level setting meetings with distinct groups of panelists, a pilot study, and an operational meeting. The purpose of the pilot study is to conduct a full “dress rehearsal” of the operational meeting, including but not limited to: an opportunity to test out materials, training procedures, collection of panelist judgments, feedback given to panelists through the process, software used to conduct analyses, meeting logistics, and other essential elements of the process. The pilot study may result in minor changes to the procedures, as well as major changes that would need additional study before being implemented in an operational meeting. The pilot study provides an opportunity for procedural validity evidence and to improve the operational meeting. At the discretion of the Governing Board, other smaller-scale studies may be conducted prior to the pilot study or in response to issues raised by the pilot study. The criteria in Principle 2a apply to panelists of both meetings.
- i) The Governing Board shall ensure that a Technical Advisory Committee on Standard Setting (TACSS) is convened to provide technical advice on all achievement level setting activities. Technical advice provided by standard setting experts throughout the project is intended to ensure that all procedures, materials, and reports are carried out in accordance with current best practices, providing additional validity evidence for the process and results. The Board or its contractor may also seek technical advice from other groups as appropriate, including NCES and the larger measurement community (e.g., the National Council on Measurement in Education).
- j) All aspects of the procedures shall have documentation as evidence of the appropriateness of the procedures and results. This evidence shall be made available to the Board at the time of deliberations about the achievement levels. A summary of the evidence shall be available to the public when the achievement level results are reported.

- k) Sample items and student responses known as exemplars shall be chosen from the pool of released items for the current NAEP assessment to reflect performance in the *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced* regions of the scale. The use of exemplars is intended to help the public better understand what students who are in each achievement level actually know and are able to do for each subject and grade. When possible, exemplars may also be chosen that reflect performance at threshold scores. The collection of exemplars shall reflect the content found in the achievement level descriptions and the range of item formats on the assessment.
- l) The outcomes from the achievement level setting panel meetings (recommended cut scores, exemplars, and ALDs for use in reporting) shall be forwarded to the Board for their consideration.

### **Principle 3: Validation and Reporting of Achievement Level Results**

**The achievement level setting process shall produce results that have validity evidence for the intended uses and interpretations and are informative to policy makers, educators, and the public.**

- a) Professional testing standards require evidence to support the intended interpretations and uses of test scores. Among the sources of evidence supporting the validity of test scores is evidence bearing on the standard setting process and results. Although standard setting is necessarily judgmental with no “true” or “correct” cut scores, the Board shall examine and consider available evidence about the procedural integrity of the achievement level setting process, the reasonableness of results, and other evidence in order to support intended uses and interpretations.
- b) NAEP achievement levels are intended to estimate the percentage of students (overall and for selected student groups) in each achievement level category, for the nation, and for states and trial urban districts (TUDAs) for some assessments. NAEP is prohibited by law from reporting any results for individual students or schools, so achievement levels do not apply to individual students or schools.
- c) To facilitate valid uses of ALDs for reporting, the Board shall ensure that the descriptions of performance for the achievement levels reflect what the empirical data reveal about the knowledge and skills of students in that score range. The Board shall revisit and may revise content ALDs following the achievement level setting to ensure that they are consistent with empirical evidence of student performance. In particular, when content ALDs are reported with results, they shall be written to incorporate empirical data from student performance. They shall describe what students at each level *do* know and *can* do rather than what they *should* know and *should* be able to do.
- d) The Board shall examine and consider all evidence related to validity of the achievement level setting activities. These data shall include, but not be limited to: procedural evidence such as training, materials and panelist evaluation data;

reliability evidence such as consistency across panelist type, subpanels, rounds, and meetings, if appropriate; and external comparisons to other similar assessments, if appropriate, with necessary caveats. The results from validation efforts shall be made available to the Board in a timely manner so that the Board has access to as much validation data as possible as it considers the recommendations regarding the final levels.

- e) In describing student performance using the achievement levels, terms such as “students performing at the *NAEP Basic* level” or “students performing at the *NAEP Proficient* level” are preferred over “*Basic* students” or “*Proficient* students”. The former implies that students have mastery of particular content represented by the achievement levels, while the latter implies an inherent characteristic of individual students.
- f) In reporting the results of NAEP, the three achievement levels of *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced* refer to the three regions of the NAEP scale at and above each respective cut score. The remaining region that falls below the *NAEP Basic* cut score shall be identified as “below *NAEP Basic*” when a descriptor is necessary.
- g) In describing the *NAEP Proficient* level, reports shall emphasize that the policy definition is not intended to reflect “grade level” performance expectations, which are typically defined normatively and can vary widely by state and over time. *NAEP Proficient* may convey a different meaning from other uses of the term “proficient” in common terminology or in reference to other assessments.
- h) An interpretative guide shall accompany NAEP reports, including specific examples of appropriate and inappropriate interpretations and uses of the results.

#### **Principle 4: Periodic Review of Achievement Levels**

**Periodic reviews of existing achievement levels shall determine whether new achievement level descriptions and/or cut scores are needed to continue valid and reliable measurement of current student performance and trends over time.**

- a) At least once every 10 years or 3 administrations of an assessment, whichever comes later, the Governing Board, through its Committee on Standards, Design and Methodology (COSDAM), shall review the alignment between the content ALDs and items, based on empirical data from past and recent administrations of NAEP assessments. In its review, COSDAM (in consultation with the Assessment Development Committee) shall solicit input from technical and subject matter experts to determine whether changes to the content ALDs are warranted or whether a new standard setting shall be conducted, making clear the potential risk of changing cut scores to trends and assessment of educational progress. Relevant factors may include but not be limited to: substantive changes in the item types or in the balance of item types; changes in the mode of administering assessments; advances in standard setting methodologies; and changes in the policy environment for using NAEP results.

- b) Within the period for a review of achievement level descriptions and cut scores, changes may occur to a NAEP framework. If a framework is replaced or revised for a major update, a new achievement level setting process may be implemented, except in circumstances where scale score trends are maintained. In this latter instance, COSDAM shall determine how to revise the ALDs and review the cut scores to ensure that they remain reasonable and meaningful.
- c) If there are major updates to a NAEP framework, the ALDs shall be updated by the Framework Visioning and Development Panel. (See the Governing Board Policy on Framework Development for additional details). Following an assessment administration under the revised framework, COSDAM may decide to use empirical data to revise content ALDs to align with the revised framework.
- d) As additional validation evidence becomes available, the Board shall review it and make a determination about whether the achievement levels should be reviewed and potentially redone.

## **Principle 5: Stakeholder Input**

**The process of developing student achievement levels is a widely inclusive activity. The Governing Board shall provide opportunities to engage multiple stakeholders throughout the achievement level setting process and shall strive to maximize transparency of the process.**

- a) The process of seeking nominations for the achievement level setting panels shall include outreach to relevant constituencies, such as: state and local educators; curriculum specialists; business representatives; and professional associations in a given content area.
- b) The Design Document (describing in detail all planned procedures for the project) shall be distributed for review by a broad constituency and shall be disseminated in sufficient time to allow for a thoughtful response from those who wish to provide one. All interested stakeholders shall have an opportunity to provide public comment.
- c) Achievement level setting panelists shall include teachers, non-teacher educators, and other interested members of the general public with relevant educational background and experience, including parents, researchers, and employers. Each panel shall reflect diversity in terms of gender, race/ethnicity, region of the country, urbanicity, and experience with students with disabilities and English language learners.
- d) All achievement level setting activities shall be informed by technical advice throughout the process. The Technical Advisory Committee on Standard Setting shall provide ongoing technical input from standard setting and assessment experts, and other groups with relevant technical expertise may be consulted periodically as needed.

- e) Ongoing input and coordination with staff and contractors from the National Center for Education Statistics (NCES) is necessary to ensure that all achievement level setting activities are carried out in a manner that is consistent with the design, analysis, and reporting of NAEP assessments.
- f) The Governing Board may ask its standing groups representing various constituencies to provide input on the achievement level setting process.

## **Principle 6: Role of the Governing Board**

**The Governing Board, through its Committee on Standards, Design and Methodology (COSDAM), shall monitor the development and review of student achievement levels to ensure that the final Governing Board-adopted achievement level descriptions, cut scores, and exemplars comply with this policy.**

- a) The Committee on Standards, Design and Methodology (COSDAM) shall be responsible for monitoring the development and review of achievement levels that result in recommendations to the Governing Board for any NAEP assessment under consideration. COSDAM shall provide direction to the achievement level setting contractor, via Governing Board staff. This guidance shall ensure compliance with the NAEP legislation, Governing Board policies, Department of Education and government-wide regulations, and requirements of the contract(s) used to implement the achievement level setting project.
- b) If there is a need to revise the initial achievement level descriptions (ALDs) created at the time of framework development for use in achievement level setting and/or reporting, the Governing Board shall take final action on revised ALDs.
- c) COSDAM shall receive regular reports on the progress of achievement level setting projects.
- d) COSDAM shall review and formally approve the Design Document that describes all planned procedures for an achievement level setting project.
- e) At the conclusion of the achievement level setting project, the Governing Board shall take final action on the recommended cut scores, exemplars, and ALDs for use in reporting. The Governing Board shall make the final determination on the NAEP achievement levels. In addition to the panel recommendations, the Board may consider other pertinent information to assess reasonableness of the results, such as comparisons to other similar assessments.
- f) Following adoption by the Governing Board, the final ALDs, cut scores, and exemplars shall be provided to the National Center for Education Statistics (NCES) for reporting the results of the NAEP assessment(s) under consideration.
- g) Consistent with Principle 4 above, COSDAM shall periodically review existing achievement levels to determine whether it is necessary to revise achievement level descriptions or conduct a new standard setting.

# **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.