Adopted: November 17, 2018

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).[1] 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.

[1] According to current NAEP legislation, the Governing Board shall develop achievement levels for all NAEP assessments except for the Long-Term Trend assessment.

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" (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
Curriculum Experts
Content Experts
Assessment Specialists
State Administrators
Local School Administrators
Policymakers
Business Representatives
Parents
Users of Assessment Data
Researchers and Technical Experts
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 (4th 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 know and 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. Each 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.

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 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 try 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 by 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 performance in each achievement level represents 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. Standard setting is necessarily judgmental, and 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) 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.

c) 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.

d) 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.

e) 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.

f) 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.

g) To facilitate valid uses of ALDs for the purpose of 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 demonstrated by students in that score range. To develop ALDs for reporting, following the achievement level setting the Board shall revisit and may revise content ALDs to ensure that they are consistent with empirical evidence of student performance. In particular, these "Reporting ALDs" chosen to illustrate the knowledge and skills demonstrated at different achievement levels shall be written to incorporate empirical data from student performance. Reporting ALDs shall describe what students at each level *do* know and *can* do rather than what they *should* know and *should* be able to do.

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 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 shall 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 revised.

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.

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 achievement level descriptions, cut scores, and exemplars recommended to the Governing Board foradoption 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 based on recommendations from COSDAM.

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 relevant 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.