

Statement on the Intended Meaning of NAEP Results (SV #3)

Background

Over the past couple of years, the Committee on Standards, Design and Methodology (COSDAM) has engaged in several discussions about the need for an explicit statement about the intended purpose and meaning of NAEP. The lack of such a statement is inconsistent with the *Standards for Educational and Psychological Testing*, which states: “The test developer should set forth clearly how test scores are intended to be interpreted and consequently used...” (AERA/APA/NCME, 2014, Standard 1.1, p. 23). The recent evaluation of NAEP achievement levels (National Academies of Sciences, Engineering, and Medicine, 2017) noted that the 2009 evaluation of the entire NAEP program recommended, “a transparent, organized validity framework beginning with a clear definition of the intended and unintended uses of NAEP assessment scores” (p. 213) and that there was no evidence that this had been implemented.

Such a statement is necessary because it gives validation efforts a “north star,” a foundation upon which to organize the considerable technical evidence that the National Center for Education Statistics (NCES) and the Governing Board already have for the defensible interpretation and use of NAEP results. It should be posted on NAEP websites and serve as the preamble for technical reports that lay out the validity evidence for NAEP. The primary purpose of this document is to provide guidance for the NAEP program and existing users of NAEP data to orient validation efforts. This is distinct from an interpretative guide for broader and more diverse audiences, which would be more engaging and include details and examples about what to do and not do with the NAEP achievement level results (and possibly scale scores as well).

COSDAM reviewed and discussed this draft statement during the spring and summer of 2019. Based on these discussions and additional follow up with Commissioner Lynn Woodworth and Associate Commissioner Peggy Carr, there was consensus that it is important to keep this statement on the purpose of NAEP to one page. Other documents that serve distinct but related purposes, such as an interpretative guide, would likely need to be longer.

During the November 2019 Board meeting, COSDAM Chair Andrew Ho provided a brief overview of the proposed statement for full Board discussion; additional edits have been incorporated to address the feedback received. Board action is planned at the upcoming meeting.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education [AERA, APA, & NCME]. (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.

National Academies of Science, Engineering, and Medicine. (2017). *Evaluation of the Achievement Levels for Mathematics and Reading on the National Assessment of Educational Progress*. Washington, DC: The National Academies Press.

The Intended Meaning of NAEP Results

The primary purpose of the National Assessment of Educational Progress (NAEP), also known as the Nation's Report Card, is to measure the educational achievement and progress of the nation's students at established grades and ages in relation to the content of NAEP frameworks. NAEP results also enable comparisons of what representative students know and can do among states and jurisdictions, among various demographic groups, and over time.

The authorizing legislation for NAEP and the National Assessment Governing Board states that the purpose of the NAEP program is broadly to, "conduct a national assessment and collect and report assessment data, including achievement data trends, in a valid and reliable manner on student academic achievement in public and private elementary schools and secondary schools..." (Public Law 107-279, Section 303(b)(2)(B)). That legislation also prohibits NAEP from maintaining any system of personally identifiable information. Thus, NAEP assesses the educational progress of groups of representative students, not individuals.

The National Center for Education Statistics (NCES) develops numerical score scales for each NAEP subject. NAEP scale scores convey the degree to which students have mastered the content described in the NAEP assessment frameworks, with higher scores indicating greater levels of mastery.

The Governing Board is charged with setting NAEP achievement levels and has established [general policy definitions](#) for *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced*. Percentages at or above achievement level cut scores indicate the percentage of students in a group who *meet or exceed* the knowledge and skills represented by specific content achievement level descriptions. These specific descriptions are found in the NAEP [assessment frameworks](#) and [reports](#). Additional information about the NAEP achievement level descriptions can be found in the Governing Board policy statement on [Developing Student Achievement Levels for NAEP](#).

NAEP results describe educational achievement for groups of students at a single point in time, progress in educational achievement for groups of students over time, and differential educational achievement and progress among jurisdictions and subpopulations.

There are several features of NAEP that distinguish it from many other assessment programs. For example:

- 1) NAEP produces results for the nation and participating states and districts. NAEP does not produce results for individual students or schools.
- 2) NAEP measures progress based on successive cohorts of students. NAEP does not produce results about the growth of individual students or groups of students over time.
- 3) NAEP results measure achievement and progress; however, NAEP results alone cannot indicate either why or how progress has occurred. Educational policies and practices that concur with NAEP progress may have caused this progress or been coincidental.
- 4) NAEP assessments are based on independent assessment frameworks developed through a national consensus approach described [here](#). NAEP frameworks do not represent any single state or local curricula.