



NAEP Achievement Level Descriptor Review Study

Executive Summary for the National Assessment Governing Board May 2022 Quarterly Meeting

May 13, 2022

Achievement Levels Overview and Study Purpose

The National Assessment Governing Board has a legislatively mandated responsibility to develop NAEP achievement levels. The current [Board policy for Developing Student Achievement Levels for NAEP](#) provides policy definitions of *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced* – and describes the principles for setting achievement levels. The policy definitions are general, high-level expectations of what students should know and be able to do and consistent across all NAEP assessments. Content achievement level descriptors (ALDs) are specific descriptions of what students at each level *should* know and be able to do for each individual assessment included in the frameworks (for example, see page 71 of the current [Mathematics Framework](#)). The [Achievement Levels Procedures Manual](#) further describes details for implementing the Board policy.

The Board first established the achievement levels policy in 1990 with the expectation that, in addition to scale scores, reporting should include the percentage of test takers at each defined level and those falling below the NAEP Basic level. As part of the NAEP reauthorization in 1994, Congress stipulated the achievement levels be designated as trial until an evaluation determined that the achievement levels are reasonable, reliable, valid, and informative to the public. The removal of the trial status is at the discretion of the NCES Commissioner.

In 2016, the National Academies of Sciences, Engineering, and Medicine conducted the [Evaluation of the Achievement Levels for Mathematics and Reading on NAEP](#), in response to the trial status stipulation. In it, they acknowledged the value of the NAEP achievement levels: “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”. They made recommendations to enhance their utility, including:

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.

In response to the recommendations presented by the National Academies and updated guidance on achievement level setting, the Board updated its achievement level policy in November 2018 with guidance to develop new achievement level descriptions of what students *can* do based on student NAEP performance. These descriptions would be used in reporting to help increase the utility of NAEP data and are referred to as Reporting ALDs. The Board approved an [Achievement Levels Work Plan](#) in 2020 to describe plans for addressing the recommendations in the evaluation.

In September 2020, the Board awarded a contract to Pearson to address the National Academies recommendation and updated Board policy for mathematics and reading. This study used the 2019 reading and mathematics NAEP assessments at grades 4, 8, and 12. This work is based on the existing mathematics and reading frameworks, not the updated versions that will be used for the 2026 administration and beyond. Once the updated frameworks are operationalized, new ALD studies will be conducted. In their report the National Academies indicated that the Board's efforts to address this recommendation could be sufficient for the removal of the trial status (depending on the outcome). The NCES Commissioner, however, indicated that additional validity evidence would be required for this to occur. Thus, removal of the trial status was not an immediate intended outcome of the current ALD study.

The primary outcomes of this study were a) the development of Reporting ALDs based on assessment items and data, and b) comparison of the Reporting ALDs to the content ALDs as validity evidence. The methodology used was specified in the Achievement Levels Work Plan and was similar to what was done to evaluate the alignment and revise the 2009 NAEP Reading ALDs for grades 4, 8, and 12 ([Donahue, Pitoniak, & Beaulieu, 2010](#)) and the 2009 NAEP Mathematics ALDs for grade 12 ([Pitoniak, Dion, & Garber, 2010](#)). COSDAM has been overseeing this work from the onset, and a Technical Advisory Committee (TAC) with six experts in achievement levels and ALDs have participated in about 100 hours of discussions to provide technical guidance on all phases of the work.

Study Process and Outcomes

The study involved convening panels of teachers and non-teacher educators with content expertise in reading or mathematics to review items, develop summary statements indicating what students know and can do as evidenced by correctly responding to the items, and then compare the statements to the existing content ALDs and provide alignment judgments.

Workshops were held virtually due to the pandemic for five full days. First, a pilot workshop was held in October 2021 to try out all procedures and identify areas needing improvement. Adjustments were made following the workshop based on this experience and recommendations from the TAC and COSDAM. Next, an operational meeting was held in February 2022 with a different set of panelists than the pilot. The operational meeting resulted in two sets of outcomes – draft reporting ALDs and final alignment judgments of reporting ALDs to the achievement level policy definitions and to the content ALDs (sent under separate cover). The reporting ALDs and alignment judgment rating results are confidential and require further review before public release. If Board members are interested in additional details about the study please contact staff to request a copy of the full report.

Next Steps

An immediate next step is to discuss and finalize the reporting ALDs for inclusion in the 2022 Report Card for reading and mathematics grades 4 and 8 (grade 12 reading and mathematics were not assessed in 2022 and will next be assessed in 2024). COSDAM members reviewed paragraph and bulleted versions of the reporting ALDs and recommend moving forward with the bulleted version, which provides clearer communication of what students at each achievement

level know and can do. The bulleted version also helps distinguish the reporting ALDs from the content ALDs. The full Board will participate in an initial discussion of the reporting ALDs at the May 2022 Board meeting, with action anticipated during the August 2022 Board meeting. The May Board session will be closed to the public because results of the study have not been released yet.

In between May and August, Board staff are collecting internal feedback on the reporting ALDs from the National Center for Education Statistics (NCES); they will be reviewing to ensure there are no concerns from an operational perspective (e.g., item security). In addition, external feedback will be sought from state and district staff who are users of NAEP data to determine whether the descriptions are clear, and to ensure that users can differentiate between achievement levels. The intention of the reviews is to determine if minor edits are needed to improve clarity or resolve operational issues. The external reviewers will not have access to the secure NAEP item pools or the extensive training panelists received; it would pose a threat to the validity of the process to make substantive changes to the reporting ALDs based on external feedback at the end of the process.

In the longer-term, COSDAM will consider the implications of the alignment judgment results for validity evidence in context of other work. COSDAM members agreed the results of the alignment judgments are independent to the decision to adopt reporting ALDs; the reporting ALDs are based on real assessment data, and thus accurate depictions of what students within an achievement level have demonstrated they know and can do.