Committee on Standards, Design and Methodology

November 17, 2022
3:00 pm – 5:00 pm ET

The Mayflower Hotel - New Jersey Room (2nd Floor)
1127 Connecticut Ave NW
Washington, DC 20036
Phone: (202) 347-3000

AGENDA

3:00 – 3:05 pm	**Welcome** *Suzanne Lane, Chair*	
3:05 – 3:20 pm	**Briefing: Two-Subject Design Investigations** *Enis Dogan, NCES*	Attachment A
3:20 – 4:30 pm	**Discussion: Communicating Achievement Levels** *Suzanne Lane* *Carey Wright, Vice Chair* *Becky Dvorak, Assistant Director for Psychometrics*	Attachment B
4:30 – 5:00 pm	**Update: Achievement Level Description Review Pilot Study for U.S. History, Civics, and Science (CLOSED)** *Eric Moyer, Pearson*	Sent under separate cover

Two-Subject Design Investigations for NAEP

The National Center for Education Statistics (NCES) recently explored the possibility of coupling a two-subject design by 2028 for grade 4 and 8 Reading and Mathematics assessments. The two-subject design would have resulted in a single student assessed in NAEP Reading and Mathematics in one administration. The expectation was that assessing each sampled student in two subject areas instead of one would permit a decrease in the required sample size, resulting in cost savings, and allow estimation of correlation between performance on these two subjects. NCES convened a working group that met periodically over several months to consider feasibility and assess the risks and benefits of a two-subject design. This session of the COSDAM meeting will describe the investigations conducted by the working group that ultimately led to the NCES decision to not pursue plans for a two-subject design for Reading and Mathematics.

Achievement Levels Communications Considerations

The Committee on Standards, Design, and Methodology (COSDAM) will continue a discussion focused on enhancing communications regarding the NAEP achievement levels as it pertains to the Achievement Levels Work Plan, adopted by the Board in 2020. In August 2022, COSDAM members identified and prioritized various stakeholder groups to target achievement level communications and considered the type and amount of information appropriate to each. They also began to consider ideas for convening an advisory Group to assist with communication efforts. Based on the discussion, Suzanne Lane (COSDAM Chair) identified next steps to assist members in further defining and achieving its goals:

- Develop an achievement levels communication planning table.
- Define the purposes of convening an achievement levels communications advisory group.
- Summarize existing achievement levels validity evidence.
- Hold informal discussions with the CCSSO and/or TUDA Task Forces to understand how they use achievement level data.

The COSDAM Chair, Vice Chair, and staff liaison took measures to accomplish each of these steps to inform the November discussion on achievement levels communications, summarized below.

In preparation for the release of 2022 NAEP Mathematics and Reading data, Board staff and communications contractors developed a four-page informational document to describe the recently developed achievement level descriptions (ALDs) for reporting. In addition to the information presented below, this document may be useful for conceptualizing future achievement level (AL) communications documents.

Following the November meeting, the next step will be to collaborate with the Reporting and Dissemination Committee (R&D) to incorporate their perspectives and develop an implementation plan.

Achievement Levels Communications Planning Table

An Achievement Levels Communications Planning table was developed based on COSDAM discussions during the August 2022 meeting and subsequent feedback on a first draft via email in September and October of 2022. The intent of the planning table is to identify stakeholders to target communications and to describe the intent of communications by stakeholder type. Color coding indicates the priority level for developing communications material based on who COSDAM expects is most likely to need and use the information (green = high; yellow = moderate; gray = low). The November meeting will be an opportunity for members to weigh in on the elements of this table before collaborating with the Board's R&D committee on achievement levels communications efforts.

Table 1. Achievement Levels Communication Planning Table

Stakeholder Type	What do we want them to know?	What do we want them to do with this information?
State School Chief/Superintendent	• General definition of NAEP • Definition and proper uses of ALs/ALDs • Policy definitions of ALDs • Explanation of national, state, and TUDA results • Example of ALDs for reporting • Brief summary of validity evidence • Availability of AL data for key subgroups Included in an extended supplemental document: • Explanation of different types/purposes of ALDs (item development versus reporting) • Link to ALDs in NAEP frameworks • More detailed validity evidence	• Inform education policy (e.g., evaluate rigor of state ALs) • Identify differences in AL performance for key subgroups and consider policy implications • Have a general idea of what students in their state in each achievement level know and can do and its implications for policy and practice • Explain NAEP ALs to others in their state, including how ALs can be used to understand student knowledge and skills • Understand and explain the difference between interpreting ALs on NAEP and state assessments • Identify initiatives needed to address areas of low performance • Communicate national and state results and explain the differences

Stakeholder Type	What do we want them to know?	What do we want them to do with this information?
State Testing Directors	- General definition of NAEP - Definition and proper use of ALs/ALDs (incl. how ALs differ from state ALs) - Explanation of national, state, and TUDA results - How and why one might link and/or compare state and NAEP ALs - Policy definitions of ALDs Example of ALDs for reporting - High-level summary of validity evidence Included in an extended supplemental document: - Explanation of different types/purposes of ALDs (item development versus reporting) - Link to ALDs in NAEP frameworks - Detailed validity evidence (e.g., standard setting information, linking studies, ALD study) - Examples of studies linking NAEP ALs to state ALs	- Examine differences in percent students (overall and by subgroup) reaching each AL on their state assessment compared to NAEP and consider implications - Consider NAEP ALs when setting their state performance levels - Compare to patterns seen in state data (overall, by subgroup) to identify differences and similarities - Provide guidance to others in their department on how to interpret NAEP AL data - Understand and explain the difference between interpreting ALs on NAEP and state assessments
Superintendents of Large School Districts	- General definition of NAEP - Definition and proper use of ALs/ALDs - Explanation of national, state, and TUDA results - Example of ALD for reporting Included in an extended supplemental document: - Explanation of different types/purposes of ALDs (item development versus reporting) - Link to ALDs in NAEP frameworks - Detailed validity evidence (e.g., standard setting information, linking studies, ALD study) - Examples of studies linking NAEP ALs to state ALs	- Explain purpose and interpretations of NAEP and NAEP ALs to district staff and school leadership - Have a general idea of what students in their state and district (if a TUDA) know and can do on NAEP assessments by achievement level - Understand how subgroups fall across the ALs - Understand and explain the difference between interpreting ALs on NAEP and state assessments

Stakeholder Type	What do we want them to know?	What do we want them to do with this information?
National and State Legislators	- General definition of NAEP - Definition and proper use of ALs/ALDs - How state ALs/ALDs differ from NAEP ALs/ALDs - Example of ALDs for reporting - Explanation of different types/purposes of ALDs	Inform educational policy, including where additional educational resources are needed (e.g., rural versus urban; student subgroups)
Education Committee Legislative Staff	- General definition of NAEP - Definition and proper use of ALs/ALDs - Policy definitions of ALDs - Example of ALDs for reporting - High-level summary of validity evidence Included in an extended supplemental document: - Explanation of different types/purposes of ALDs (item development versus reporting) - Link to ALDs in NAEP frameworks - Detailed validity evidence (e.g., standard setting information, linking studies, ALD study) - Examples of studies linking NAEP ALs to state ALs	- Inform educational policy, including where additional educational resources are needed (e.g., rural versus urban, student subgroups) - Understand how to interpret NAEP AL results and explain to legislators

Stakeholder Type	What do we want them to know?	What do we want them to do with this information?
Business Community	• General definition of NAEP • Definition and proper use of ALs/ALDs • policy definitions of ALDs • Explanation of national, state, and TUDA results • Example of ALDs for reporting • High-level summary of validity evidence Included in an extended supplemental document: • Explanation of different types/purposes of ALDs (item development versus reporting) • Link to ALDs in frameworks • Detailed validity evidence • Examples of studies linking NAEP ALs to State ALs	• Use NAEP AL data to identify states/TUDAs that may benefit from additional resources to advance students • Have a general idea of what students at each achievement level know and can do and consider implications for workforce readiness
Journalist	• General definition of NAEP • Definition and proper use of ALs/ALDs • policy definitions of ALD • Explanation of national, state, and TUDA results • Example of ALDs for Reporting • High-level summary of validity evidence Included in an extended supplemental document: • Explanation of different types/purposes of ALDs (item development versus reporting) • Link to ALDs in frameworks • Detailed validity evidence (e.g., standard setting information, linking studies, ALD study) • Examples of studies linking NAEP ALs to State ALs	• Understand and describe what NAEP is and how to interpret AL results • Describe AL results accurately and articulately for a wide audience • Pull examples from ALDs for reporting to help distinguish between what students at different ALs can likely do • Interpret and report subgroup differences in achievement level percentages

Stakeholder Type	What do we want them to know?	What do we want them to do with this information?
State School Board Members	• General definition of NAEP Definition and proper use of ALs/ALDs • Policy definitions of ALDs • Example of ALD for reporting • High-level summary of validity evidence	• Understand what NAEP is and how to interpret ALs • Inform where resources are most needed (e.g., rural versus urban, student subgroups)
Educators	• General definition of NAEP • Definition and proper use of ALs/ALDs • Policy definitions of ALDs • Example of ALD for reporting • High-level summary of validity evidence	• Understand what NAEP is and how to use it to examine student achievement at the national and state levels • Understand the difference between interpreting ALs on NAEP and state assessments
Parents and Students	• General definition of NAEP • Definition and proper use of ALs/ALDs • Example of ALDs for reporting	• Understand what NAEP is and how to use it to examine student achievement at the national and state levels • Understand the difference between interpreting ALs on NAEP and state assessments
General Public	• General definition of NAEP • Definition and proper use of ALs/ALDs • Example of ALD for reporting	• Understand what NAEP is and how to use it to examine student achievement at the national and state levels • Understand the difference between interpreting ALs on NAEP and state assessments
State Test Developers and Content Standards Developers	• General definition of NAEP • Definition and proper use of ALs/ALDs • How and why one might link and/or compare state and NAEP ALs • Policy definitions of ALDs • Example of ALD for reporting • High-level summary of validity evidence	• Examine differences in percent students reaching each AL on their state assessment compared to NAEP and consider implications • Examine what students at each AL know and can do on NAEP and compare with what students are expected to know and do by the state • Understand similarities and differences between the knowledge and skills their state assessment vs NAEP and consider implications for their state performance levels

Considerations for an Achievement Levels Communications Advisory Group

The Achievement Levels Work Plan calls for convening an ongoing advisory group to discuss and provide feedback on the development of materials for communicating NAEP achievement levels. Some examples of how the advisory group might be used include a) providing input on the different stakeholder groups (or audiences) who would benefit from receiving communication materials and for what purposes; b) providing input on content for inclusion in communication materials; c) reviewing materials to ensure messaging is clear and concise to desired audiences while maintaining technical accuracy; and d) assisting the Board/NCES in determining if additional efforts are advisable to make informed decisions (e.g., focus groups, white papers). COSDAM members began to brainstorm potential composition of the advisory group in August 2022. Initial recommendations include individuals with strong communication and reporting skills, testing expertise, education task force and/or advocacy members, and business leaders. Efforts to conceptualize and implement an advisory group will be forthcoming and will include collaboration with additional Board members (e.g., R&D) and NCES staff.

Validity Evidence

The Achievement Levels Work Plans calls for compilation and documentation of validity evidence to support intended interpretations and uses of NAEP achievement levels. The table below presents existing validity evidence to consider in validity evidence documentation.

Table 2. Sources of achievement levels validity evidence.

Document Type	Validity Evidence
Achievement Levels Policy	The Board's Achievement Levels Policy describes the procedures followed by the Board for developing and reviewing achievement levels. The policy calls for rigorous, research-based procedures, providing evidence that procedural validity is a priority.
Standard setting technical documentation	Standard setting reports present the procedures used for setting the cut points associated with each NAEP achievement level on the NAEP assessments. These reports illustrate the use of rigorous, research-based procedures, thus providing procedural validity evidence.
Linking studies reports: NAEP and SAT for FL NAEP and PIRLS NAEP and ACT Explore for KY, NC, and TN NAEP and SAT for MA NAEP and ACT nationally and for TN and MI NAEP and postsecondary preparedness for MI	Various linking studies have been conducted that demonstrate how NAEP achievement levels correspond to external academic measures. Particularly, studies provide insight into how the achievement levels correspond to performance on college entrance exams and postsecondary performance. These studies provide external validity evidence of NAEP achievement levels.
Achievement Level Description Review Study for Reading and Mathematics technical report (available December 2022)	The Achievement Level Description Review study summarizes content included on the NAEP Mathematics and Reading assessments and compares this content with what was intended, as presented in the content ALDs included in the framework. This work provides content validity evidence of the achievement levels.

Summary of CCSSO Task Force Achievement Level Discussion

CCSSO Task Force members participated in a discussion led by Becky Dvorak (Assistant Director for Psychometrics) and Suzanne Lane, COSDAM Chair, on October 12, 2022. The purpose of this discussion was to understand how members use NAEP data, achievement level data specifically, in their states. This information is intended to inform ongoing COSDAM activities related to achievement level communications and utility. The Task Force members represented 12 state departments of education: Alabama, Arizona, Connecticut, Florida, Kentucky, Nebraska, Oregon, Texas, Virginia, and Wisconsin. Task Force members were asked the following questions:

1. What do you focus on when you review NAEP results for your state?
   a. Do you focus on scale scores, achievement levels, or both?
   b. If you examine achievement levels – what aspects of the achievement levels do you focus on? (e.g., all levels, the percentage reaching *NAEP Proficient*, how results have changed over time, subgroup level results)
   c. If you do not examine the achievement levels – why don't you find them useful? What could be done so that they would be useful?
   d. Have you spent time investigating what it means to reach the different NAEP Achievement Levels, that is, what knowledge and skills students need to reach the achievement levels? If so, how? (e.g., reading achievement level descriptions, reviewing the NAEP item map).
2. How have you used NAEP results, in particular the achievement levels, to inform decisions in your state?
3. How could the achievement levels be improved to be more useful to you?

Multiple CCSSO task force members reported that they used NAEP results to examine their rank against other states, particularly their rank against other states in their region. Members who responded to this question indicated the scale scores were the first thing they examined.

Achievement levels were examined by members as well, particularly the percent of students reaching *NAEP Proficient*. One member noted they examined the percent reaching *NAEP Proficient* over time. Multiple members expressed NAEP achievement level data were used to validate their state achievement levels. One noted that after they updated their state standards to be more rigorous and fewer students reached state proficiency, they were able to support the increased rigor by comparing against NAEP data. Two members expressed NAEP achievement level data were used to inform state standard settings. One CCSSO leader in attendance noted that their state attempted to match the state achievement levels to NAEP psychometrically, with some degree of success.

Some concerns were raised regarding NAEP achievement level interpretations. For example, one Task Force member indicated it was difficult to understand that only 30% of students reach *NAEP Proficient,* while a higher percentage attend and graduate college. Multiple members expressed confusion with the different interpretations of *NAEP Proficient* versus proficiency on state assessments. One member noted that before the No Child Left Behind (NCLB) legislation

states focused on passing rates, not achievement levels. Once NCLB passed it required states to address proficiency similarly to NAEP, which led to the current confusion. The NAEP program responded by referring to its proficiency level as *NAEP Proficient* instead of only Proficient; however, some Task Force members felt this distinction is not always sufficient for differentiation.

In addition to scale scores and achievement levels, one member reported utility in reporting NAEP data by percentiles. When they saw increases for high performers and decreases for low performers on state assessment data, they were able to look to NAEP data broken down by percentiles as a source of validation. A second member echoed the utility of percentile data.

When asked to consider what could make NAEP data more useful, one Task Force member suggested identifying the point at which students reach grade-level. This would be challenging, however, given that states differ in their definitions of grade-level. Lane asked Task Force members whether it would be useful for NAEP to define an achievement level at below the *NAEP Basic* cut. Two members thought this could be useful because of the increase in students falling in the below *NAEP Basic* range. Another member expressed equity concerns in using the label "below" to describe students' performance.

Task members also discussed the utility of the new reporting achievement level descriptions (ALDs) developed for NAEP Mathematics and Reading. These ALDs provide a summary of what students demonstrated they likely know and can do based on NAEP performance. One task member thought the ALDs would be helpful for making achievement levels more useful and for understanding the requirements of NAEP versus those of states. Another member felt the reporting ALDs would most likely be used by education advocates and researchers, but probably not by the average reporter or user. Another member indicated they found the item maps more useful than the reporting ALDs for understanding what students can do within an achievement level. One Task Member chimed in that the item maps and reporting ALDs are likely both useful for different situations.