



# COMMITTEE ON STANDARDS, DESIGN AND METHODOLOGY (COSDAM) AGENDA

---

**May 16, 2024**

**4:00 – 6:00pm EST**

**Space 2**

---

**4:00 – 4:05 pm**

**Welcome**

*Suzanne Lane, Chair*

---

**4:05 – 4:35 pm**

**NAEP Achievement Levels Activities:  
Interpretations and Intended Audiences**

*Suzanne Lane*

Attachment A

---

**4:35 – 4:55 pm**

**COSDAM Discussion of Effect Sizes and  
Practical Significance**

*Suzanne Lane*

*Rebecca Dvorak, Assistant Director for  
Psychometrics*

Attachment B

---

**4:55 – 5:00 pm**

***Transition to Joint Committee Meeting in  
A.Boardroom***

---

**5:00 – 6:00 pm**

**Joint Meeting with Reporting and  
Dissemination (R&D) Committee:  
Considering an Interpretive Guide**

*Suzanne Lane, Chair, COSDAM*

*Marty West, Chair, Reporting and  
Dissemination Committee*

Attachment C

---

## **NAEP Achievement Levels Activities: Interpretations and Intended Audiences**

**May 16, 2024**

### **Goal**

This session is intended to a) present the language on appropriate and inappropriate uses of achievement levels being proposed for inclusion in the NAEP Achievement Levels Validity Argument, currently under development, for COSDAM input, and b) remind members of the intended audiences for various achievement levels materials that should be kept in mind during development of the validity argument, and of communications strategies that will be discussed during a joint meeting with the Reporting and Dissemination (R&D) committee following the COSDAM meeting.

### **Overview**

This session will include an overview of the current status of the NAEP Achievement Levels Validity Argument, currently under development, and plans for finalization. The committee Chair Suzanne Lane will present the proposed language on appropriate and inappropriate uses and interpretations of NAEP Achievement Levels planned for the concluding section of the report based on the available evidence and will allow time for discussion if there are any additional considerations.

### **Background**

A NAEP Achievement Levels Validity Argument is currently under development and a full draft for COSDAM review is expected to be delivered the summer of 2024. This validity argument was called for in the [Achievement Levels Work Plan](#), adopted by the Board in 2020 to address recommendations presented by the National Academies of Sciences, Engineering, and Medicine (NASEM) in their 2017 evaluation of the NAEP Achievement Levels.

COSDAM members have offered ongoing input into the development of the report, including in March 2024 when the feedback primarily focused on how to best target achievement level information to various audiences. The validity argument is intended to be a comprehensive synthesis of validity evidence and though it will be made publicly available, it will be best suited for those with some knowledge of assessment – including measurement experts and staff from state departments of education. However, COSDAM expects some of the information will be useful to pull for inclusion in communications strategies intended for a wider audience – particularly information regarding appropriate and inappropriate uses and interpretations of NAEP Achievement Levels.

COSDAM discussion led to initial input into the language on appropriate and inappropriate uses and interpretations for inclusion in the validity argument, presented

in Tables 1 and 2 below. In addition, based on COSDAM feedback and information compiled during the development of the validity argument thus far, the contractor developing the argument has been offered the following guidance:

- As much as possible, make this section of the validity report accessible to a wide audience compared to previous sections. This is meant to draw conclusions from the evidence, and may not require highly technical language compared to prior sections.
- Clarify the purposes and uses of NAEP Achievement Levels versus state achievement levels that are important for understanding how interpretations between the assessments differ.
- Based on COSDAM feedback, tying in how publicly available item maps and released NAEP items can help aid in understanding of what it means to perform at each NAEP Achievement Level.

**Table 1. Draft of appropriate interpretations and uses of NAEP Achievement Levels**

<b><u>Appropriate Uses of NAEP Achievement Levels (ALs)</u></b>	<b>Possible Sources of Evidence</b>
Performance at <i>NAEP Proficient</i> represents a solid understanding of subject-matter content	Policy and technical documentation of AL development and the standard setting process
Though not directly related to state achievement levels, NAEP ALs can help inform the comparisons of state achievement level cut-points	State Mapping Studies; information on AL development and the standard setting process; State achievement level documentation
Reporting Achievement Level Descriptions (ALDs) provide information on what students performing at each AL can likely do based on assessment data <ul style="list-style-type: none"> <li>• Include full set of Reporting ALDs for validity argument document; one or two examples for briefs.</li> </ul>	ALD Study reports for Reading and Mathematics, and for U.S. History, Civics, and Science
AL performance is related to other/academic and college readiness outcomes <ul style="list-style-type: none"> <li>• Performance at higher NAEP ALs in grade 12 associated with greater likelihood of attending a two- or four- year college</li> <li>• Performance in <i>NAEP Advanced</i> associated with a greater likelihood of majoring in a STEM field in college compared to other achievement levels</li> </ul>	Linking study reports, including: <ul style="list-style-type: none"> <li>• NAEP linked with High School Longitudinal Study of 2009 (HSL:09)</li> <li>• NAEP linked with Early Childhood Longitudinal Study (ECLS-K:2011)</li> <li>• Various studies linking NAEP with college entrance exams</li> </ul>

<b><u>Appropriate</u> Uses of NAEP Achievement Levels (ALs)</b>	<b>Possible Sources of Evidence</b>
<ul style="list-style-type: none"> <li>Performance at <i>NAEP Proficient</i> or above in grade 4 Reading associated with higher reading trajectories in elementary school</li> </ul>	

**Table 2. Draft of inappropriate interpretations and uses of NAEP Achievement Levels**

<b><u>Inappropriate</u> Uses of NAEP Achievement Levels</b>	<b>Possible Sources of Evidence</b>
Using NAEP Reading ALs to determine the percentage of students that can or cannot read	Policy and technical documentation of AL development and standard setting process, information from framework and Reporting ALDs; State achievement level documentation
The percent <i>NAEP Proficient</i> (or <i>NAEP Basic</i> , or <i>NAEP Advanced</i> ) indicates the percentage of students falling at grade level for a given subject	Policy and technical documentation of AL development and standard setting process; note regarding how NAEP achievement levels differ from state achievement levels; State achievement level documentation
Using NAEP AL data as an outcome measure to determine cause and effect impacts of state- or district-level interventions	Policy and technical documentation of AL development; Intended Meaning of NAEP Results; External information on requirements for determining causality

In March of 2023 COSDAM participated in a joint meeting with the Reporting and Dissemination (R&D) Committee to consider audiences for targeting communications strategies that would include information about how to best interpret NAEP Achievement Levels. R&D and COSDAM members agreed that prioritizations should be made to communicate with journalists, state department of education and district assessment staff, and those in education policy making positions (e.g., governors, state legislators on education committees and their staff). Though other stakeholders are important (e.g., teachers, students, parents) they are most likely to use information from state or classroom assessments that provide student- and school-level information. For more information about prior discussions regarding how to best target materials to specific stakeholders, see materials prepared for the and [May 2023](#) COSDAM meetings.

# COSDAM Discussion of Effect Sizes and Practical Significance

**May 16, 2024**

## **Goal**

This session is intended to prepare members of the Committee on Standards, Design and Methodology (COSDAM) for a joint staff meeting with the Reporting and Dissemination (R&D) committee by recalling prior discussions related to effect size and practical significance that occurred prior to many of the COSDAM members joining the National Assessment Governing Board.

## **Overview**

This session will provide an overview of prior discussions by COSDAM regarding effect size and practical significance, including the limitations with statistical significance that led to these discussions. This session will also describe and provide an opportunity to consider interpretability concerns associated with reporting effect sizes.

## **Background**

The Nation's Report Card uses significance testing to highlight score differences. If a score difference is found to be non-significant, it indicates the values should be interpreted as effectively the same. If a score difference is found to be significant, it indicates confidence that the difference represents a real difference in the full population, though it does not inform the size of the difference.

Alternatively, effect sizes provide practical meaning to score differences. They represent the size of the difference or change in scores. Effect sizes are measured in terms of how many standard deviations away from each other the two numbers are. They are computed based on the size of the difference and the variability of scores across the samples, as measured by standard deviations. Larger differences and smaller variabilities are associated with greater effect sizes – which indicate larger differences. Appendix B of the [August 2023](#) COSDAM materials include background information with examples of incorporating effect sizes with NAEP data.

Over the past year, COSDAM has discussed limitations to reporting only statistical significance to examine NAEP score differences and changes, and has considered whether it may be useful to also incorporate effect sizes. COSDAM members have acknowledged that effect sizes may be useful for researchers, psychometricians, and those with statistics backgrounds, but may be difficult to interpret to others.

# **Joint Meeting of the Committee on Standards, Design and Methodology and the Reporting and Dissemination Committee on Interpreting NAEP Results**

**May 16, 2024**

## **Goal**

At the May 2024 Governing Board meeting in Alexandria, Virginia, the Reporting and Dissemination (R&D) Committee will hold a joint meeting with the Committee on Standards, Design and Methodology (COSDAM). The purpose of this joint meeting is to discuss and develop an outline, scope, approach, and content for a guide to interpreting National Assessment of Educational Progress (NAEP) results, which the Board would make public to aid in the initial release of 2024 NAEP data.

## **Overview**

Conversations with Board members, media, stakeholders such as those at policy research organizations and advocacy groups, as well as state and district task force members point to the challenge of interpreting NAEP data with accuracy, meaning, and ease. There is widespread appreciation for the value and integrity of the Nation's Report Card and the data the National Center for Education Statistics (NCES) provides with robust, accessible visualizations. However, many have requested a guide on how to interpret the statistics and how to avoid misNAEPery—the misuse of NAEP data. Members of the two committees will convene together at this quarterly meeting to discuss a path forward on creating an interpretive guide.

## **Background**

Interest in an interpretive guide is not new, though the current scope of interest may differ from earlier conceptualizations. From 2014 to 2016, the National Academies of Sciences, Engineering, and Medicine evaluated the NAEP achievement levels in mathematics and reading, which are the responsibility of the Governing Board. In their evaluation, the National Academies' sixth recommendation urged guidance "to help users determine inferences that are best made with achievement levels and those best made with scale score statistics."

In May 2018, COSDAM and R&D met to discuss developing an "interpretative guide [which] shall accompany NAEP reports, including specific examples of appropriate and inappropriate interpretations and uses of the results" (Principle 3i).

The Academies' recommendations and this previous collaboration between R&D and COSDAM focused solely on achievement levels, but with the immense attention paid to the 2022 NAEP results, by those familiar with NAEP and by many unfamiliar with NAEP, guidance on how to interpret the statistics more broadly should prove useful.

The Board published documents intended to facilitate interpretability of NAEP Achievement Levels for the [2022 NAEP Reading and Math](#) and the [2022 NAEP U.S. History and Civics](#) releases. These documents present publicly-released items, to show example skills and knowledge that students at each achievement level likely know and can do. By connecting achievement levels to actual items, which reflect actual content, NAEP audiences can more readily distinguish the levels and interpret their meaning. These documents require revision, as they were developed quickly, and incorporate only some of the information R&D and COSDAM hope to include for the upcoming releases of 2024 NAEP data.