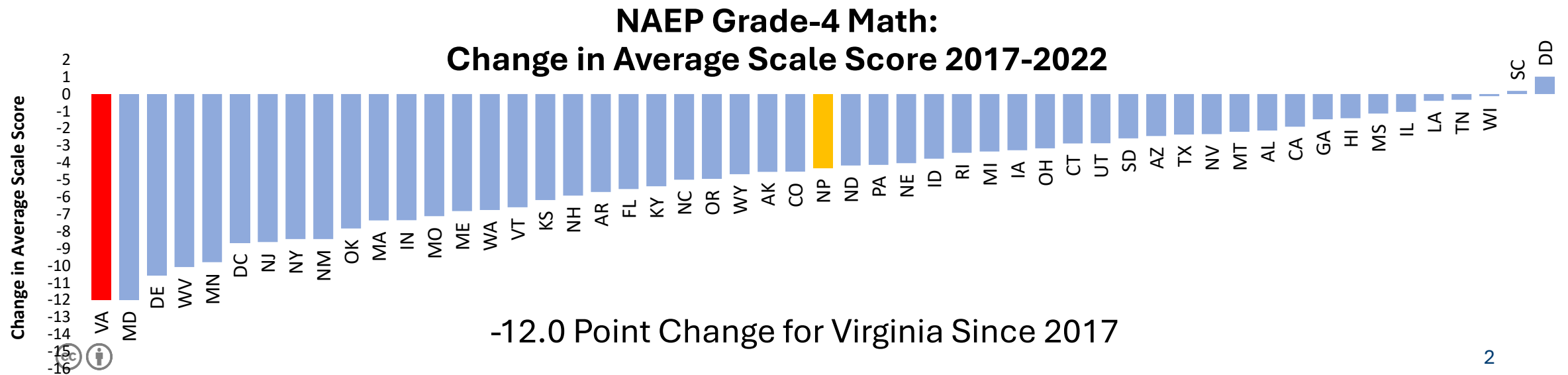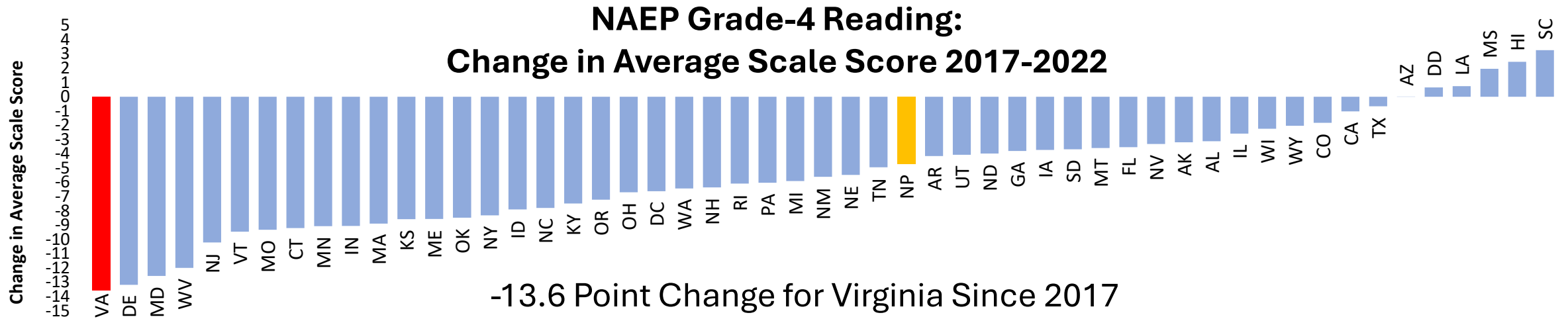# Raising the Bar: Virginia's Standard Setting Process

Virginia Board of Education

August 2024

Virginia NAEP Performance 2017-2022

NAEP Grade-4 Reading:
Change in Average Scale Score 2017-2022

-13.6 Point Change for Virginia Since 2017

NAEP Grade-4 Math:
Change in Average Scale Score 2017-2022

-12.0 Point Change for Virginia Since 2017

The Honesty Gap: Virginia Grade 4 NAEP Performance 2017-2022

Grade-4 Reading: Percent Proficient and Above

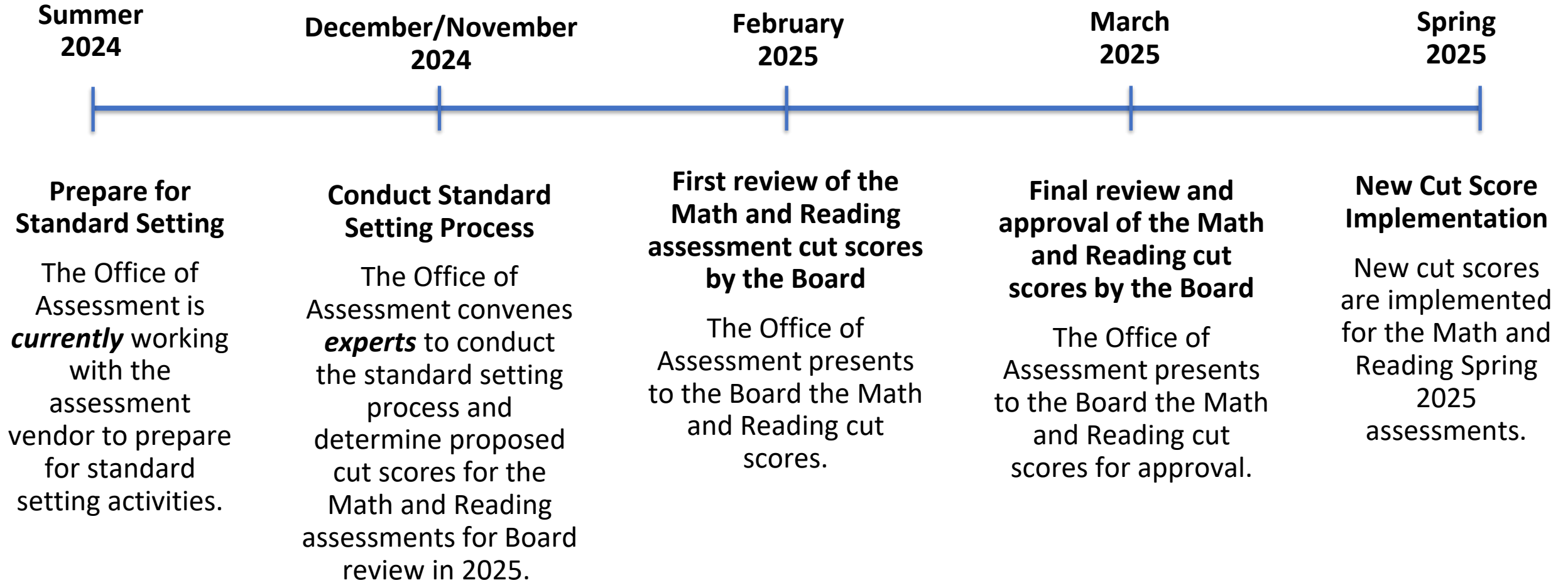Grade-4 Math: Percent Proficient and Above

■— Virginia SOL   ■— Virginia NAEP   ▬ Honesty Gap   ▬ Change in SOL Cut Scores

3

Timeline for Standard Setting

Summer 2024

December/November 2024

February 2025

March 2025

Spring 2025

Prepare for Standard Setting

The Office of Assessment is *currently* working with the assessment vendor to prepare for standard setting activities.

Conduct Standard Setting Process

The Office of Assessment convenes *experts* to conduct the standard setting process and determine proposed cut scores for the Math and Reading assessments for Board review in 2025.

First review of the Math and Reading assessment cut scores by the Board

The Office of Assessment presents to the Board the Math and Reading cut scores.

Final review and approval of the Math and Reading cut scores by the Board

The Office of Assessment presents to the Board the Math and Reading cut scores for approval.

New Cut Score Implementation

New cut scores are implemented for the Math and Reading Spring 2025 assessments.

Setting Performance Expectations for the Virginia Standards of Learning (SOL) Assessments

Chris Domaleski, Ph.D.

Associate Director

National Center for the Improvement of Educational Assessment

Chris Domaleski

- Associate Director of the Center for Assessment

- Currently working with 35+ states, districts, and other entities to provide technical and design support for a range of assessment and accountability issues

- Non-partisan and independent of any governmental agency or testing company

- I currently serve as a member of Virginia's Technical Advisory Committee

Virginia's Technical Advisory Committee

- An assessment Technical Advisory Committee, commonly referred to as a TAC, provides the Virginia Department of Education with technical assistance and support across its general assessments.

- TAC members are individuals with extensive background and experience with large scale educational assessments.

- Virginia's TAC members:
  - Greg Cizek, Ph.D. (University of North Carolina)
  - Chris Domaleski, Ph.D. (Center for Assessment)
  - Suzanne Lane, Ph.D. (University of Pittsburgh)
  - Barbara Plake, Ph.D. (retired, University of Nebraska)
  - Denny Way, Ph.D. (formerly at the College Board)

What the TAC does in Virginia

Regularly meet with the VDOE to provide guidance on topics such as:

1. Standard Setting Options

2. Test Development Strategy

3. Item Selection Process

4. Operational Protocol Feedback

What is standard setting?

- A process used to determine the minimum performance required for each performance level. Minimum performance is operationalized as a "cut score."

- Standard setting includes judgment from subject matter experts who know the test-taking population and assessed content.

- The credibility of cut scores is closely related to the <u>process</u>. It is important to design and implement an established process with a strong research base.

Virginia Performance Levels				
Grades 3-8 Mathematics and Reading SOL tests	Pass/Advanced	Pass/Proficient	Fail/Basic	Fail/Below Basic
End of Course, History & Social Science, and Science SOL tests	Pass/Advanced	Pass/Proficient	Fail/Does Not Meet	

Why does this matter to the Board?

- The Board plays a key role in approving the cut scores for state assessments that result from the standard setting process.

- Given the Board's role, it is important that the Board understand the process and the rigor/research put into the standard setting process.

What influences the rigor of the standards?

Content standards

- State content standards describe 1) what students should be able to do and 2) how well they should be able to do it

The assessment

- Assessment items are written to address the content standards. Item type and complexity matter (i.e., more innovative items designed to elicit higher-order thinking can support more rigorous performance standards).

- Form specifications and blueprints are created based on the measurable skills in the content standards. The assessment should reflect the breadth and depth of the standards

Policy Decisions

- This is often reflected in Policy Descriptors

Virginia Policy Level Descriptors

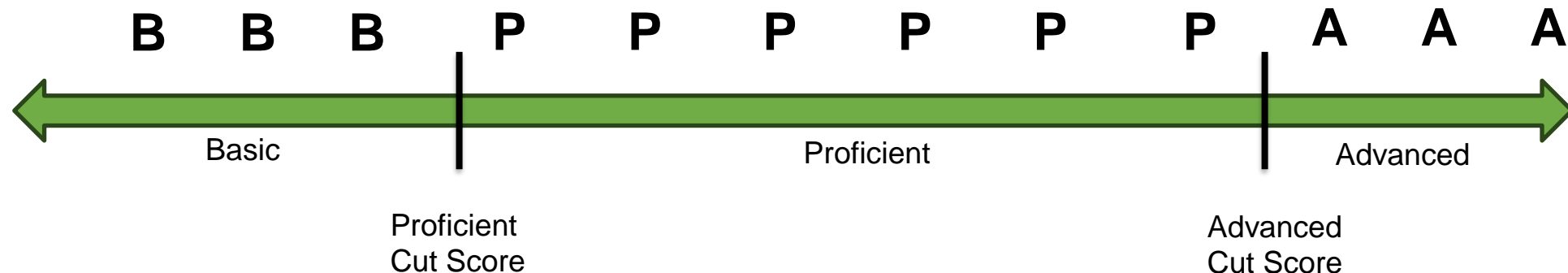| Policy Level Descriptors for Virginia Performance Levels			
Fail/Below Basic	Fail/Basic	Pass/Proficient	Pass/Advanced
The student demonstrates a limited understanding of rigorous academic knowledge aligned with the Virginia Standards of Learning for the grade/course. The student will require significant support with grade level content and skills.	The student demonstrates an inconsistent understanding of rigorous academic knowledge aligned with the Virginia Standards of Learning for the grade/course. The student will need additional support with grade level content and skills.	The student demonstrates a consistent understanding and application of rigorous academic knowledge and skills aligned with the Virginia Standards of Learning for the grade/course.	The student demonstrates an extensive depth of understanding and a consistent, accurate application of rigorous academic knowledge and skills aligned with the Virginia Standards of Learning for the grade/course.

Standard Setting

- Standard setting operationalizes the PLDs in the form of cut scores on the state assessments
- There are different methods for setting standards:
  - Item centered approaches such as Angoff, Bookmark, ID matching
  - Examinee centered approaches such as borderline groups, contrasting groups
- Item centered approaches are much more common for criterion-referenced assessments such as the Virginia SOLs
- There is not a consensus "best" method for standard setting as long as it is research-based and upholds established professional practices such as:
  - Includes well-qualified experts who follow the procedures with fidelity
  - Center's judgments in the Performance Level Descriptors
  - Includes efforts to promote coherence (i.e., 'articulation')

Virginia's Standard Setting Method: ID Matching Procedure

Center for Assessment

General procedures

- Review test and PLDs; receive training

- In multiple rounds, participants review an Ordered Item Book (OIB) and assign items to a performance level based on the extent to which the item matches the expectations in the PLDs

- Ratings are refined as participants evaluate and discuss the results and consider impact data

- Cut scores are determined by identifying the threshold region

- Examine results across grades and adjust as needed

- Document process and results, submit for review

B   B   B   P   P   P   P   P   P   A   A   A

Basic   Proficient   Advanced

Proficient
Cut Score

Advanced
Cut Score

Advantages of the ID Matching Procedure

- It is well-established, having been used by numerous large-scale state programs. Although the procedure has been around since at least 1993, it's use has grown considerably in the past decade.

- The cognitive task for participants is often viewed as more straightforward and is more directly tied to the performance standards than some other methods.

- It's a flexible method and can be adapted to incorporate the priorities of the state program.

www.nciea.org

Chris Domaleski
cdomaleski@nciea.org