New York State Education Department Assessment Technical Advisory Committee and Standard Setting

Presentation to the New York State Board of Regents

March 13, 2023

New York State EDUCATION DEPARTMENT
Knowledge › Skill › Opportunity

Introduction

Goals:

1. To introduce you to the composition and role of the New York State Education Department's Assessment Technical Advisory Committee (TAC)

2. To preview the process of setting cut scores on our state assessments

What is the Role of the Technical Advisory Committee?

The Technical Advisory Committee (TAC) is an advisory board composed of nationally known and respected psychometricians and other professionals in the field of student assessment appointed to advise the Board of Regents and the New York State Education Department (NYSED).

The TAC Advises NYSED on:

- Current and proposed policy relating to testing and test development;

- Issues related to grade-by-grade testing;

- Emerging educational and assessment issues;

- Designing research programs and how to implement them; and

- A broad-based comprehensive assessment validity agenda.

New York State TAC Co-Chairs

Howard Everson

Professor, City University of New York

Senior Research Scientist, SRI International

Marianne Perie

President, Measurement in Practice

Director, Assessment Research & Innovation, WestEd

The New York State TAC

Robert Brennan
Professor Emeritus,
University of Iowa

Derek Briggs
Professor, University of
Colorado

Claudia Flowers
Professor, University of
North Carolina, Charlotte

Andrew Ho
Professor, Harvard

Richard Luecht
Professor, UNC,
Greensboro

Suzanne Lane
Professor, University of
Pittsburgh

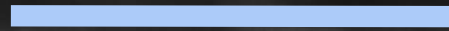**Mari Pearlman**
Owner, Pearlman Education
Group

James Pellegrino
Professor, University of
Illinois, Chicago

Michael Rodriguez
Professor, University of
Minnesota

Standard Setting

"

...cut scores embody value judgments as well as technical and empirical considerations.

—Standards in Psychological and Educational Testing, 2019

What is Standard Setting?

Applying **professional judgment** to the question "How much is *just enough*?"

It results in recommended cut scores that separate performance levels.

Types of Standards

Content Standards

- Specify the grade-level or course **content** that students should learn
  - ➢ In September 2017, the Board of Regents approved the New York State Next Generation Learning Standards for English Language Arts and Mathematics, which became effective at the beginning of the 2022-2023 school year.
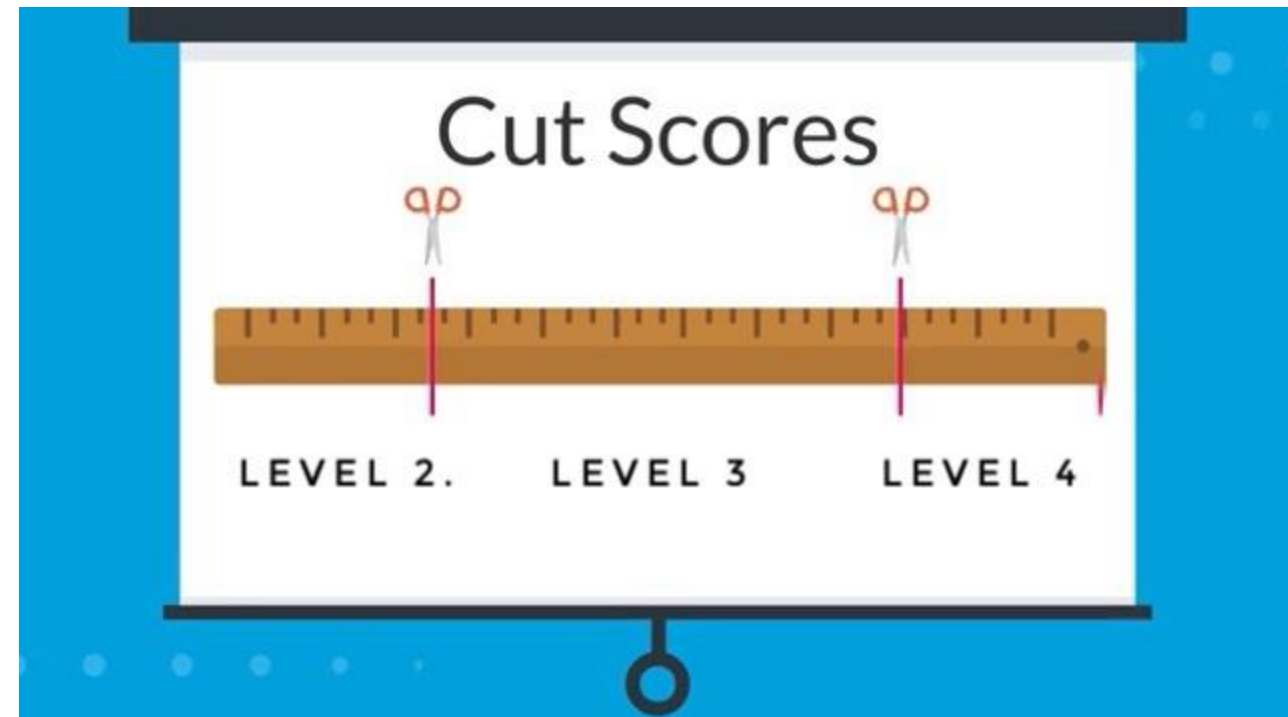
Performance Standards

- Specify the **amount** of knowledge and/or skills relative to the content standards required to achieve an outcome or classification
  - ➢ Performance Level Descriptors (PLDs)
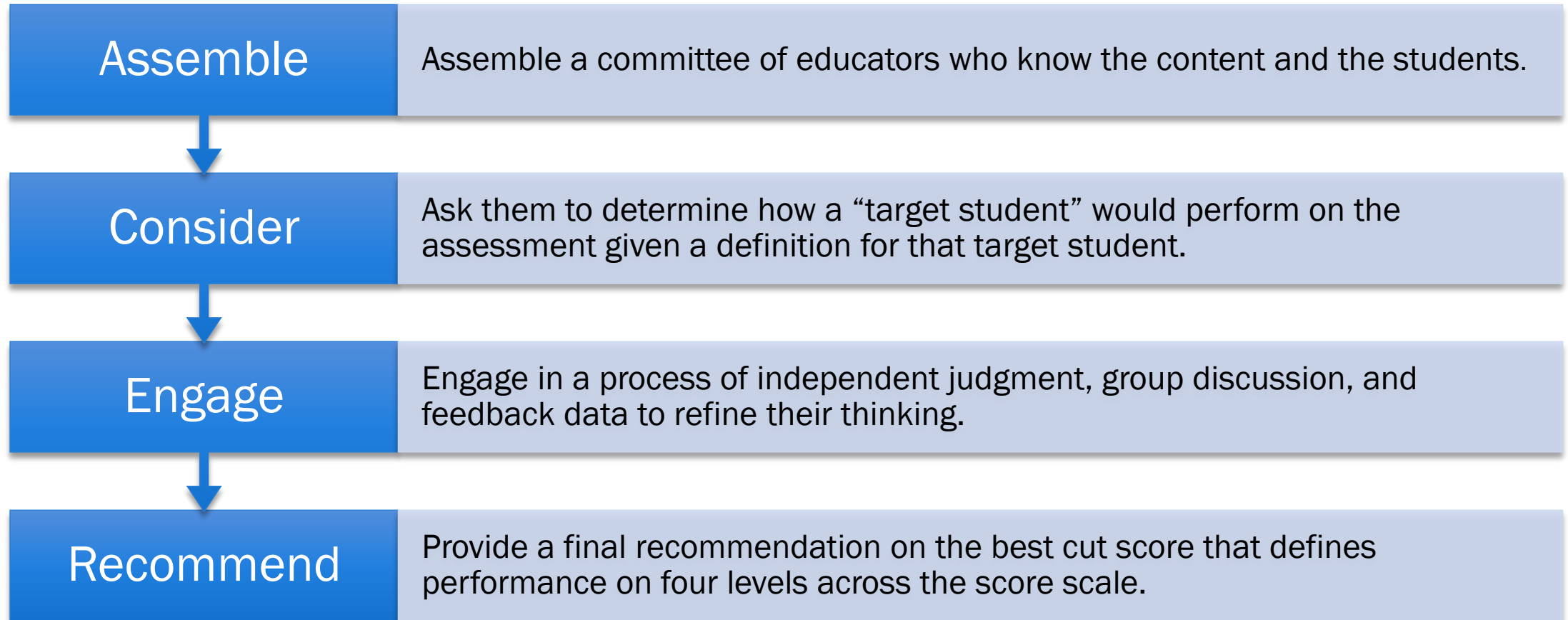 - ➢ Cut Scores
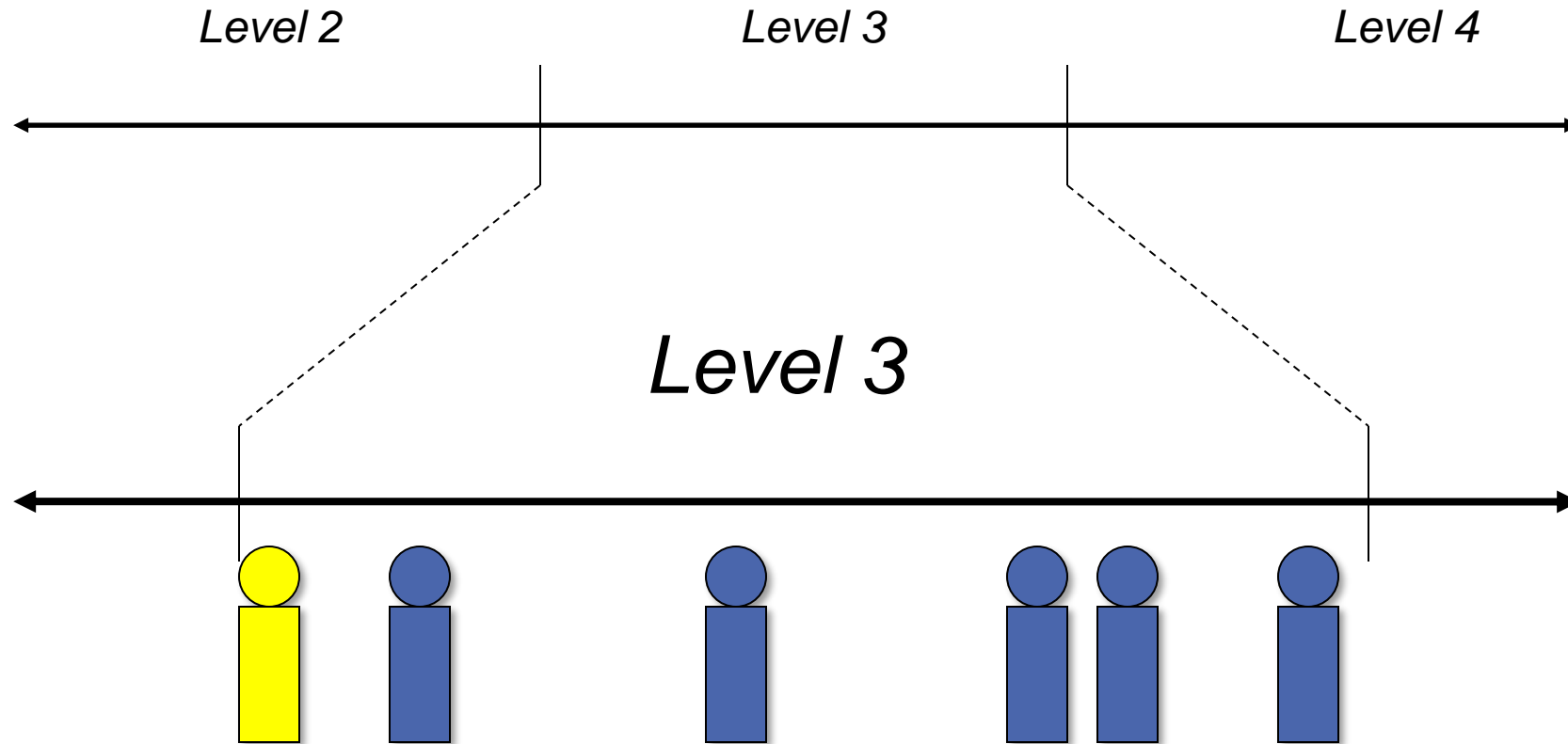
What is a Performance Standard?

Performance level descriptors (PLDs) say in words what is meant by Levels 2, 3, and 4.

What is the general design for standard setting?

Assemble	Assemble a committee of educators who know the content and the students.
Consider	Ask them to determine how a "target student" would perform on the assessment given a definition for that target student.
Engage	Engage in a process of independent judgment, group discussion, and feedback data to refine their thinking.
Recommend	Provide a final recommendation on the best cut score that defines performance on four levels across the score scale.

NYS ED .gov

Who is the target student in standard setting?

The Target Student = Borderline

The idea of borderline performance is key to any standard setting procedure.

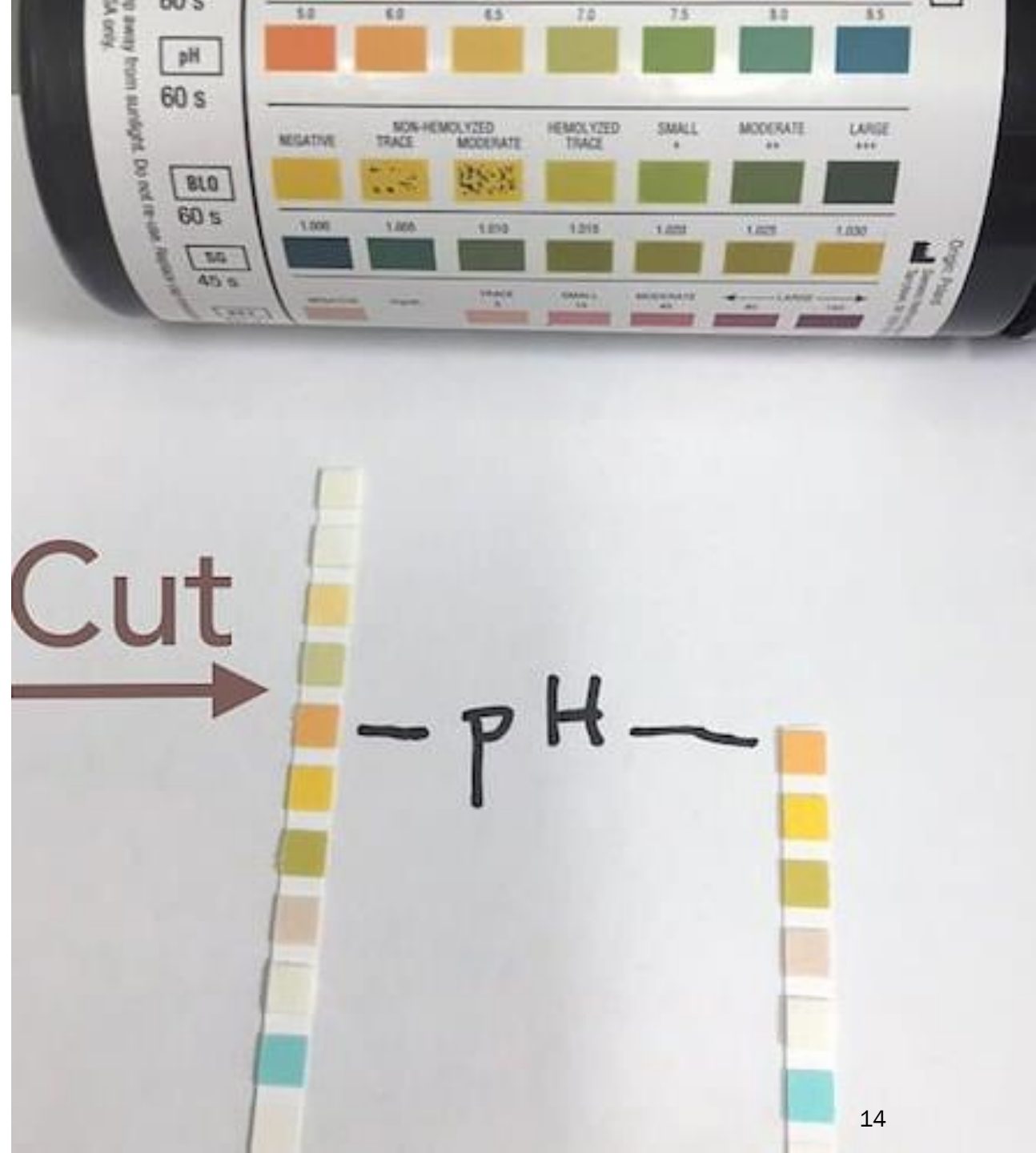Most performance level descriptions describe the entire range, not just borderline.

Because borderline performance is directly related to the placement of the cut score, a good portion of standard setting must be spent gaining a common understanding of what it means to be "borderline."

Borderlines in Education and Other Fields

Points Between:

- Highest bacteria count in safe water/
  Lowest bacteria count in polluted water

- Worst essay that deserves score of 6 /
  Best essay that deserves a 5

- Best performing "proficient" student /
  Worst performing "advanced" student

- Lowest score that passes /
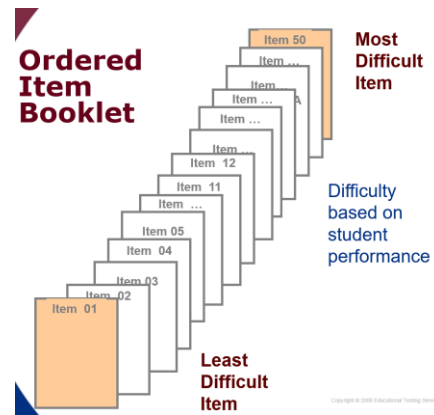 Highest score that fails

Methodology

Method

- There are multiple research-based methods for setting cut scores

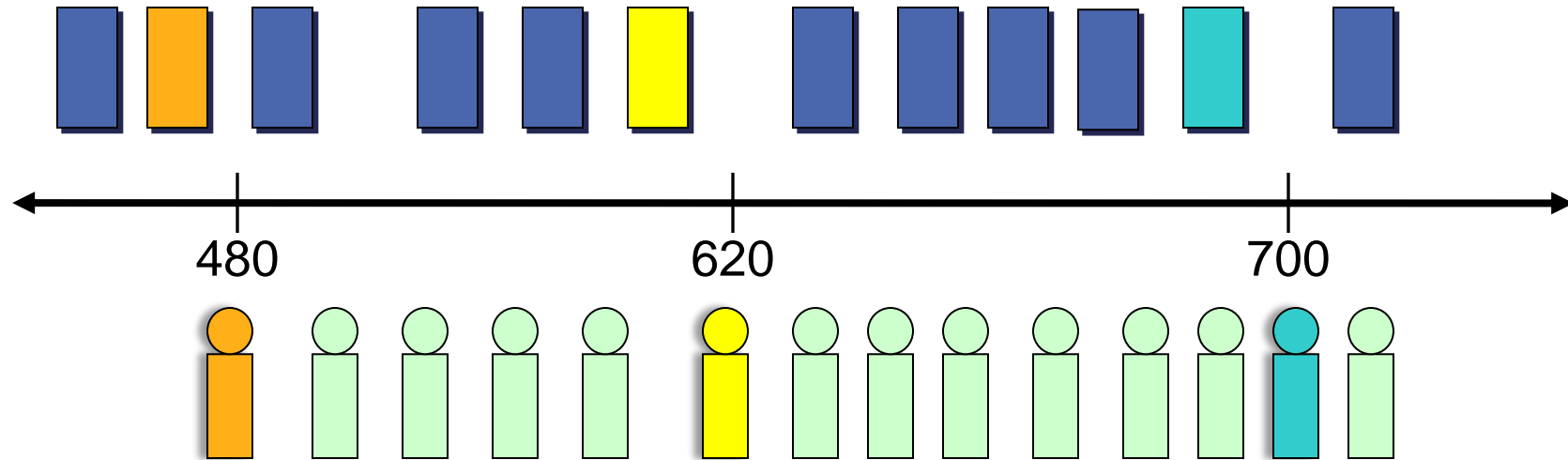- New York most frequently uses Bookmark, which is the most commonly used method across states

Tool

Process

- Use a set of items ordered by difficulty (Ordered Item Booklet, OIB)

- Find the place in the OIB that separates performance levels

- Place a bookmark at that location and record bookmark placement

Students and Items Can Be Scaled Together

The last item that a target student can answer correctly 2/3 of the time identifies the area of the cut score.

Content Expertise

This process sets criterion referenced cut scores (based on content) not normative (fixed percentage).

The educators bring their content expertise to bear to match the knowledge and skills required to answer each item to the knowledge and skills listed in the performance level descriptors.

They make judgments independently, but then talk with the other educators in their panel to hear different viewpoints.

Data Added To The Process

Feedback data

- The ordered item booklet gives them some information on item difficulty as they are ordered based on student performance.

- Panels are given information on how many students would be in each performance level based on the most recent administration of the test and their cut scores.

Other reference data can be provided

- Historical data
- Other state tests
- NAEP

Policy

Ultimately, cut scores are a policy decision.

Educators provide recommendations based on content; policymakers consider results compared to the purpose and use of the assessments.

Summary

TAC Role

Advise the State on technical issues around testing

Cut Score Process

A combination of policy and technical considerations applied through a standardized process

Policymakers' Role

Provide guidance on the policy that informs performance standards

Thank you

Howard Everson

heverson@gc.cuny.edu

Marianne Perie

mp@measurementinpractice.com