

# 2013 U.S. History EOC Assessment Standard Setting

Pam Stewart  
Commissioner  
Florida Department of Education

Vince Verges  
Director of Assessment  
Florida Department of Education

Dr. Serena Lin  
Pearson

August 13, 2013



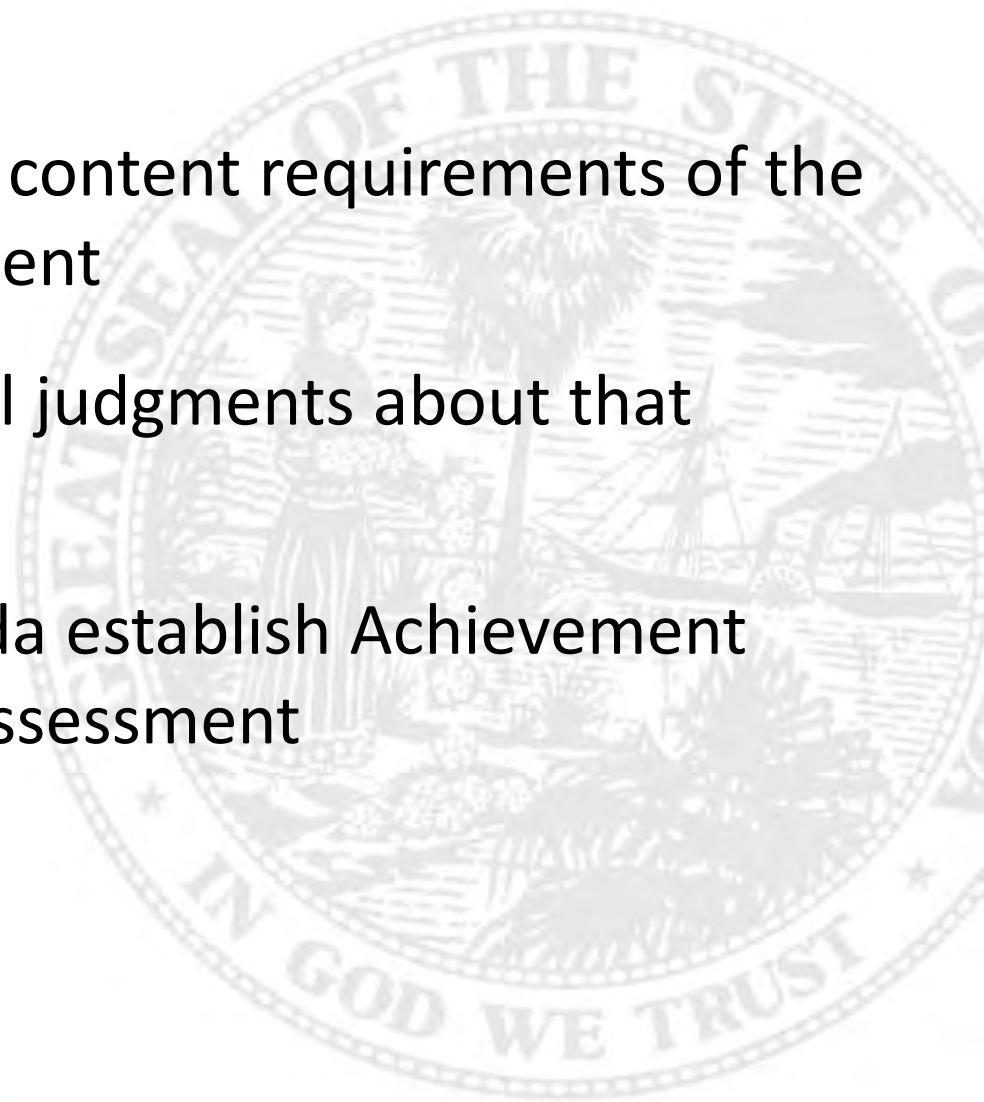
# Welcome & Introductions

- Florida Educators
- Florida Department of Education
- Pearson



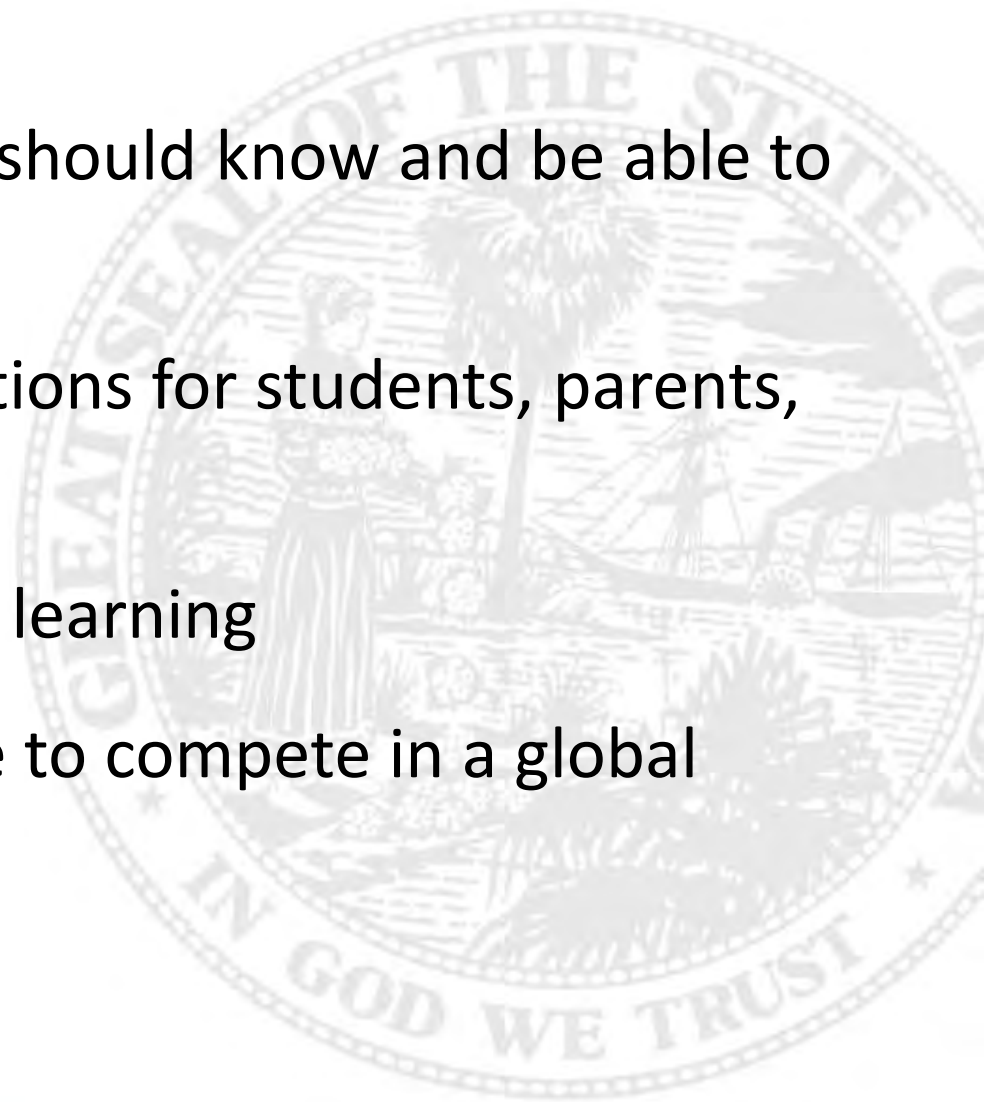
# Overview – What are you doing here?

- To thoroughly review the content requirements of the U.S. History EOC Assessment
- To provide your individual judgments about that content
- To help the State of Florida establish Achievement Level standards for this assessment



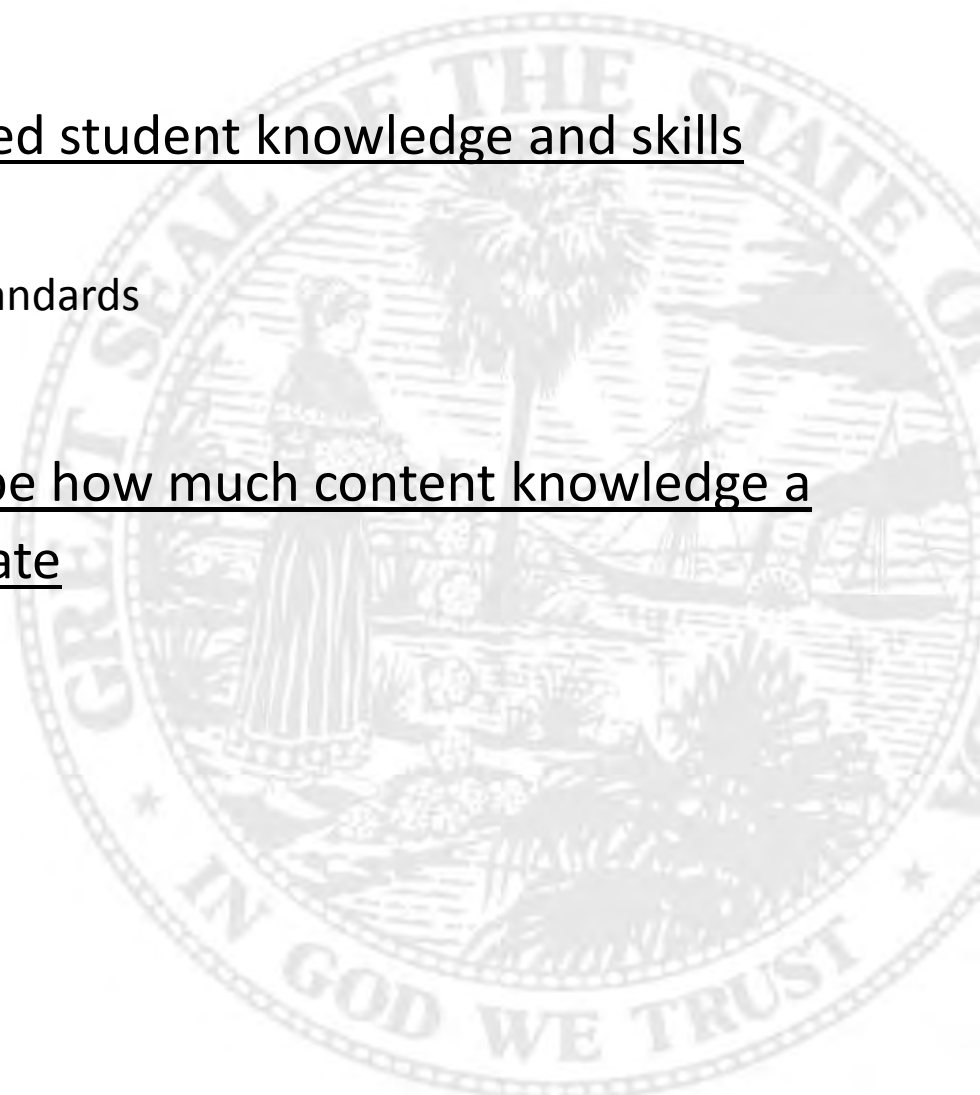
# Why Have Standards?

- To define what students should know and be able to do
- To identify clear expectations for students, parents, and teachers
- To improve teaching and learning
- To develop a society able to compete in a global economy



# Types of Standards

- Content Standards: Define desired student knowledge and skills (the “what”)
  - Next Generation Sunshine State Standards
  - Common Core State Standards
- **Performance Standards:** Describe how much content knowledge a student is required to demonstrate
  - ***Achievement Level Standards***
  - ***Graduation Requirement***
- Accountability Standards
  - School Grading Criteria
  - Annual Measurable Objectives



# Transition Schedule

Type of Assessment	Assessment Area	Year Administered to Students			
		2011-12	2012-13	2013-14	2014-15
<b>FCAT</b>	FCAT Writing	Gr 4, 8, 10			
<b>FCAT 2.0</b>	FCAT 2.0 Writing		Gr 4, 8, 10	Gr 4, 8, 10	
	FCAT 2.0 Reading	Gr 3-10	Gr 3-10	Gr 3-10	
	FCAT 2.0 Mathematics	Gr 3-8	Gr 3-8	Gr 3-8	
	FCAT 2.0 Science	Gr 5, 8	Gr 5, 8	Gr 5, 8	Gr 5, 8
<b>Florida End-of-Course Assessments</b>	Algebra 1	In Course	In Course	In Course	In Course
	Geometry	In Course	In Course	In Course	In Course
	Biology 1	In Course	In Course	In Course	In Course
	<b>U.S. History</b>		<b>In Course</b>	<b>In Course</b>	<b>In Course</b>
	Civics (Middle School)			In Course	In Course
<b>Common Core State Standards Assessments</b>	English Language Arts				Gr 3-11
	Mathematics				Gr 3-8
	High School Mathematics EOCs (Algebra 1, Geometry, Algebra 2)				In Course



# U.S. History EOC Assessment

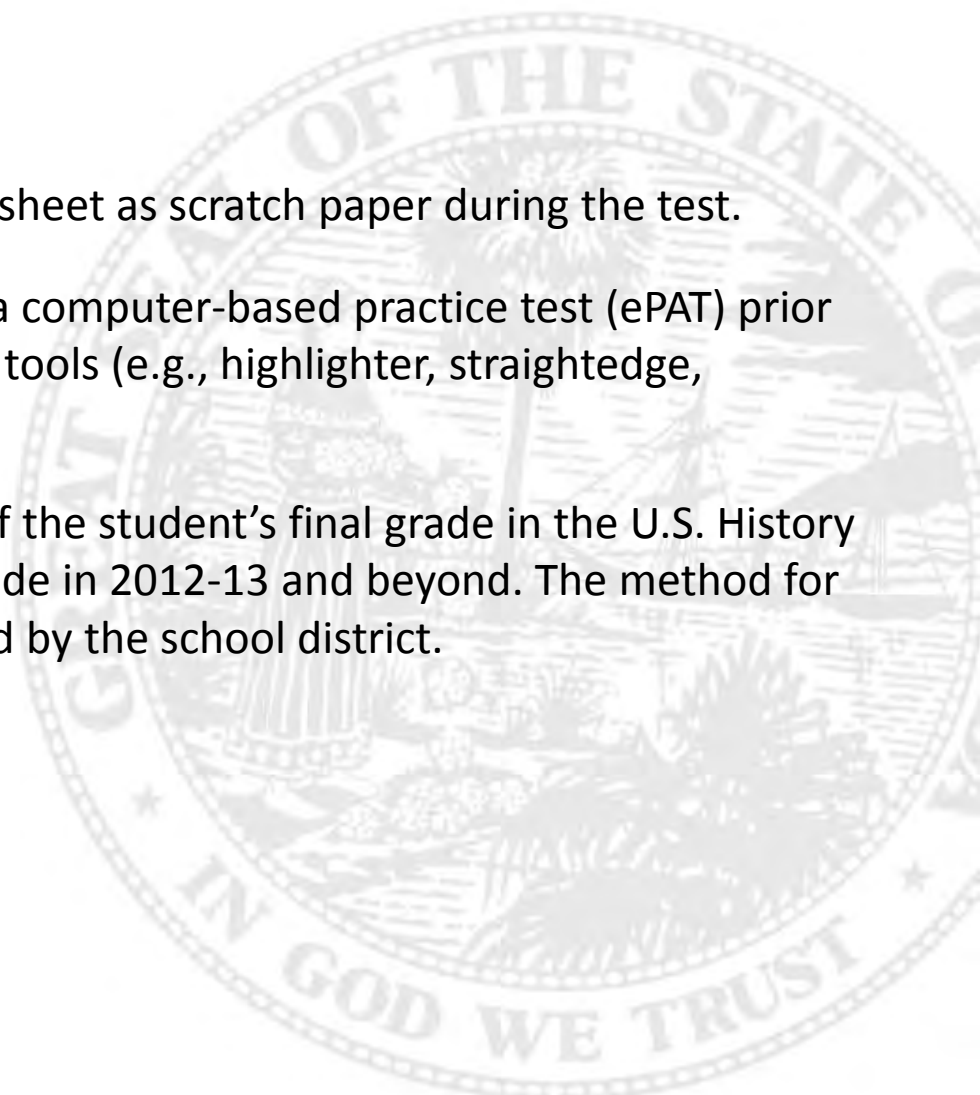
## Administration Summary

- Administered to all students enrolled in and completing one of the following courses:
  - United States History – 2100310
  - United States History Honors – 2100320
- Computer-based assessment with computer-based accommodations (e.g., screen reader, zoom, color contrast) and paper-based accommodations (e.g., regular print, large print, braille) available for students with disabilities who require allowable accommodations, as specified in their individual educational plans (IEPs) or Section 504 plans.
- Administered in one 160-minute session with a 10-minute break after the first 80 minutes. Any student not finished by the end of the allotted time may continue working, but the student must finish within the same school day.

# U.S. History EOC Assessment

## Administration Summary (continued)

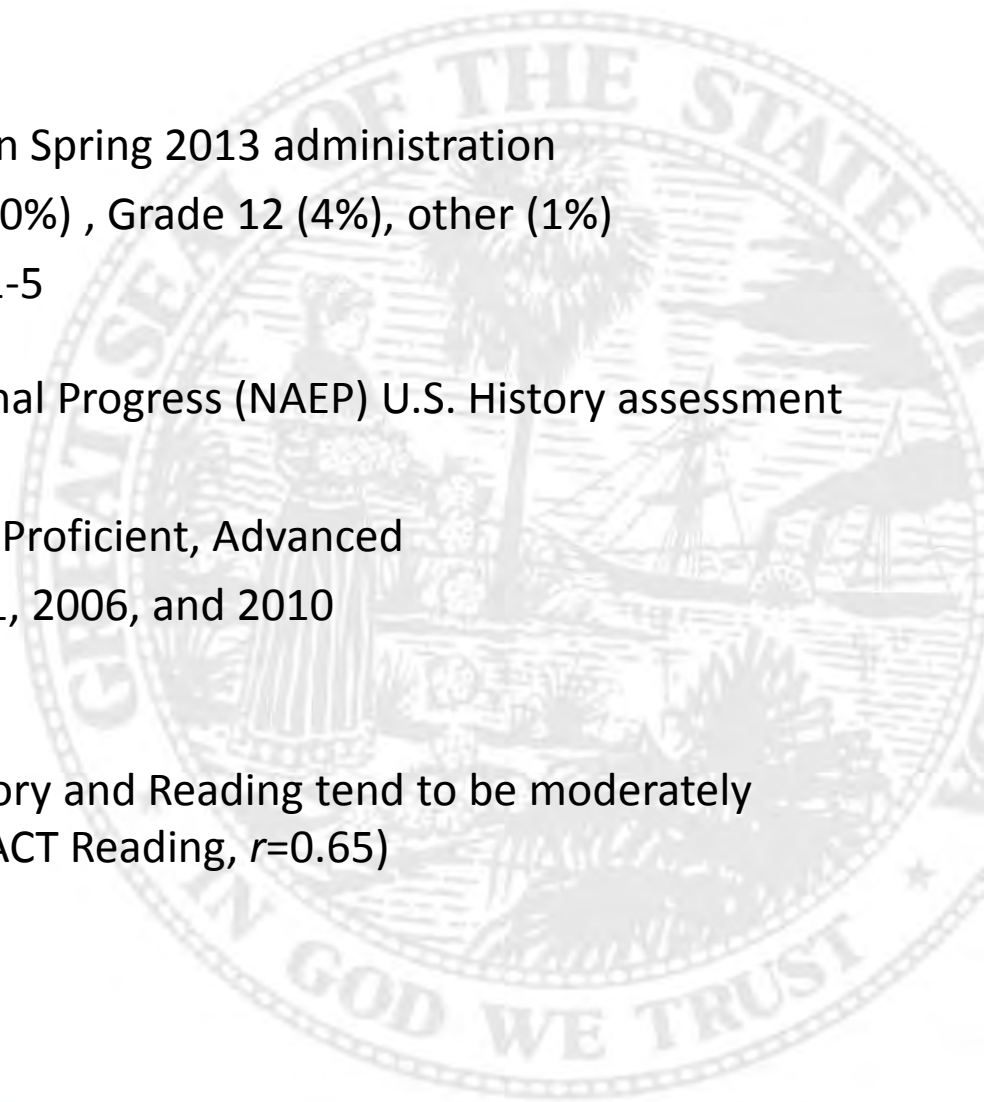
- Students may use a one-page CBT worksheet as scratch paper during the test.
- Students are required to participate in a computer-based practice test (ePAT) prior to the assessment to practice using the tools (e.g., highlighter, straightedge, notepad, option eliminator).
- Scores must be used to calculate 30% of the student's final grade in the U.S. History course for students who entered 9<sup>th</sup> grade in 2012-13 and beyond. The method for applying this requirement is determined by the school district.





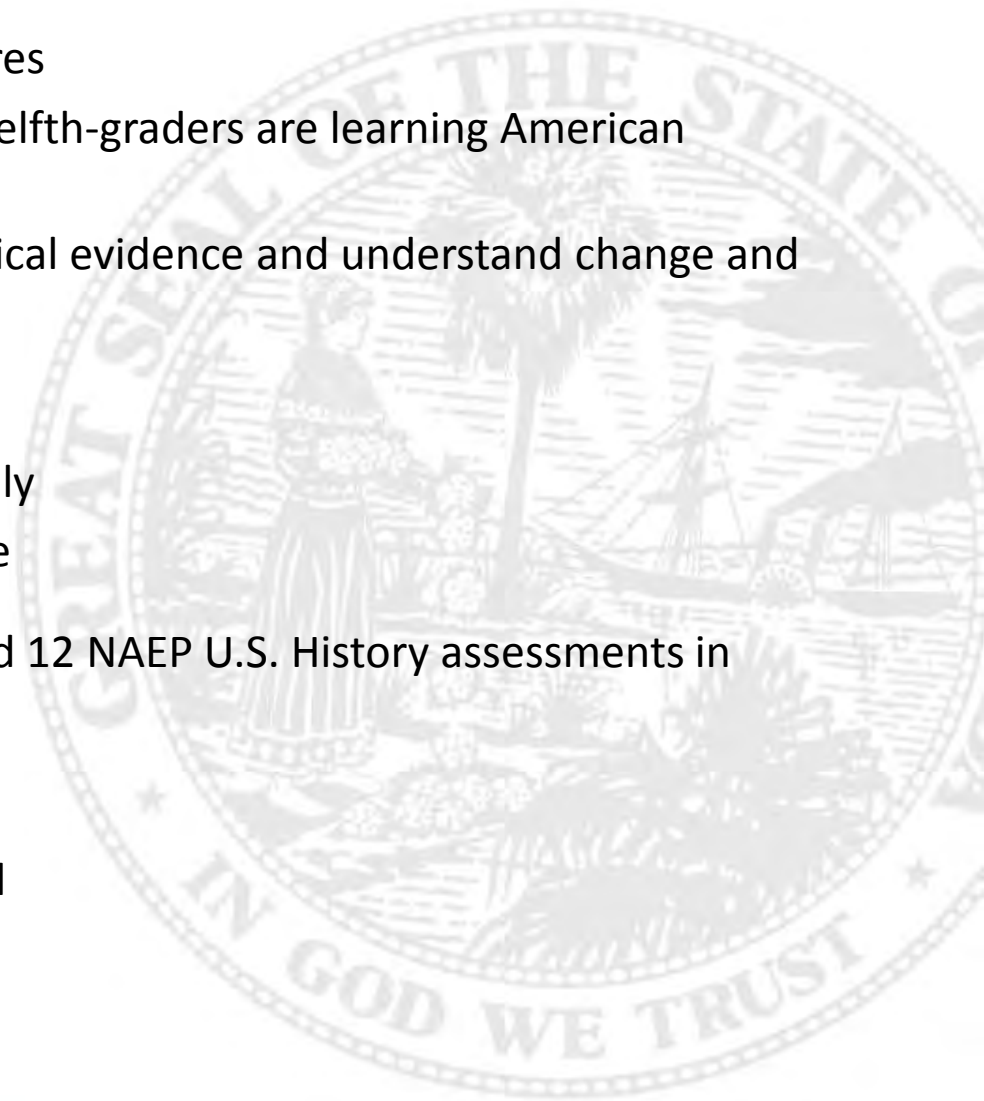
# Data to be Discussed

- FL U.S. History EOC Assessment
  - Examinee composition by grade in Spring 2013 administration
    - Grade 11 (85%), Grade 10 (10%) , Grade 12 (4%), other (1%)
  - Five Achievement Levels: Levels 1-5
  
- The National Assessment of Educational Progress (NAEP) U.S. History assessment
  - Tested in grades 4, 8, and 12
  - Has 3 Achievement Levels: Basic, Proficient, Advanced
  - Florida participated in 1994, 2001, 2006, and 2010
  
- Grade 10 FCAT 2.0 Reading
  - Student performance in U.S. History and Reading tend to be moderately correlated (STAAR U.S. History – ACT Reading,  $r=0.65$ )
  - 2011-2013 trend data



# NAEP U.S. History

- NAEP U.S. History assessment measures
  - how well fourth-, eighth-, and twelfth-graders are learning American history, and
  - whether they can evaluate historical evidence and understand change and continuity over time
- NAEP U.S. History assessment results
  - published at the national level only
  - no state-level results are available
- Florida participated in grades 4, 8, and 12 NAEP U.S. History assessments in 1994, 2001, 2006, and 2010
- Achievement Level Classification
  - NAEP: Basic, Proficient, Advanced
  - FL EOC: Levels 1, 2, 3, 4, and 5



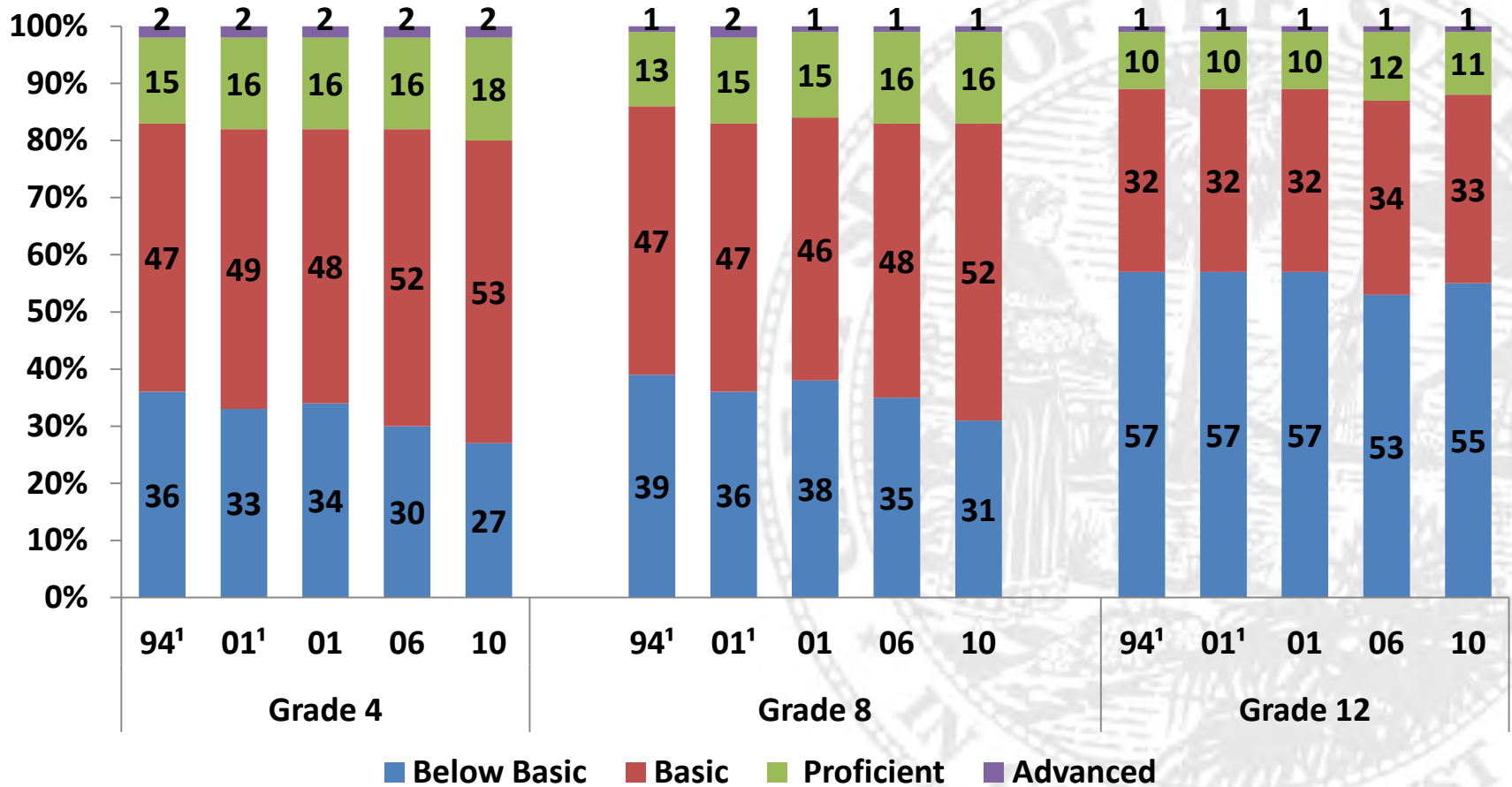
# NAEP U.S. History

## Grade 12 Achievement Level Descriptions

- |                   |   |
|-------------------|---|
| <b>Basic</b>      | <ul style="list-style-type: none"> <li>▪ Identify the significance of many people, places, events, dates, ideas, and documents in U.S. history.</li> <li>▪ Recognize the importance of unity and diversity in the social and cultural history of the United States, and have an awareness of America’s changing relationships with the rest of the world.</li> <li>▪ Have a sense of continuity and change in history and be able to relate relevant experience from the past to their understanding of contemporary issues.</li> <li>▪ Recognize that history is subject to interpretation and should understand the role of evidence in making an historical argument.</li> </ul> |
| <b>Proficient</b> | <ul style="list-style-type: none"> <li>▪ Understand particular people, places, events, ideas, and documents in historical context, with some awareness of the political, economic, geographic, social, religious, technological, and ideological factors that shape historical settings.</li> <li>▪ Communicate reasoned interpretations of past events, using historical evidence effectively to support their positions.</li> <li>▪ Their written arguments should reflect some in-depth grasp of issues and refer to both primary and secondary sources.</li> </ul>  |
| <b>Advanced</b>   | <ul style="list-style-type: none"> <li>▪ Demonstrate a comprehensive understanding of events and sources of U.S. history.</li> <li>▪ Evaluate historical claims critically in light of the evidence.</li> <li>▪ Understand that important issues and themes have been addressed differently at different times and that America’s political, social, and cultural traditions have changed over time.</li> <li>▪ Write well-reasoned arguments on complex historical topics and draw upon a wide range of sources to inform their conclusions.</li> </ul>  |

# Trends in NAEP U.S. History Achievement

## At the National Level



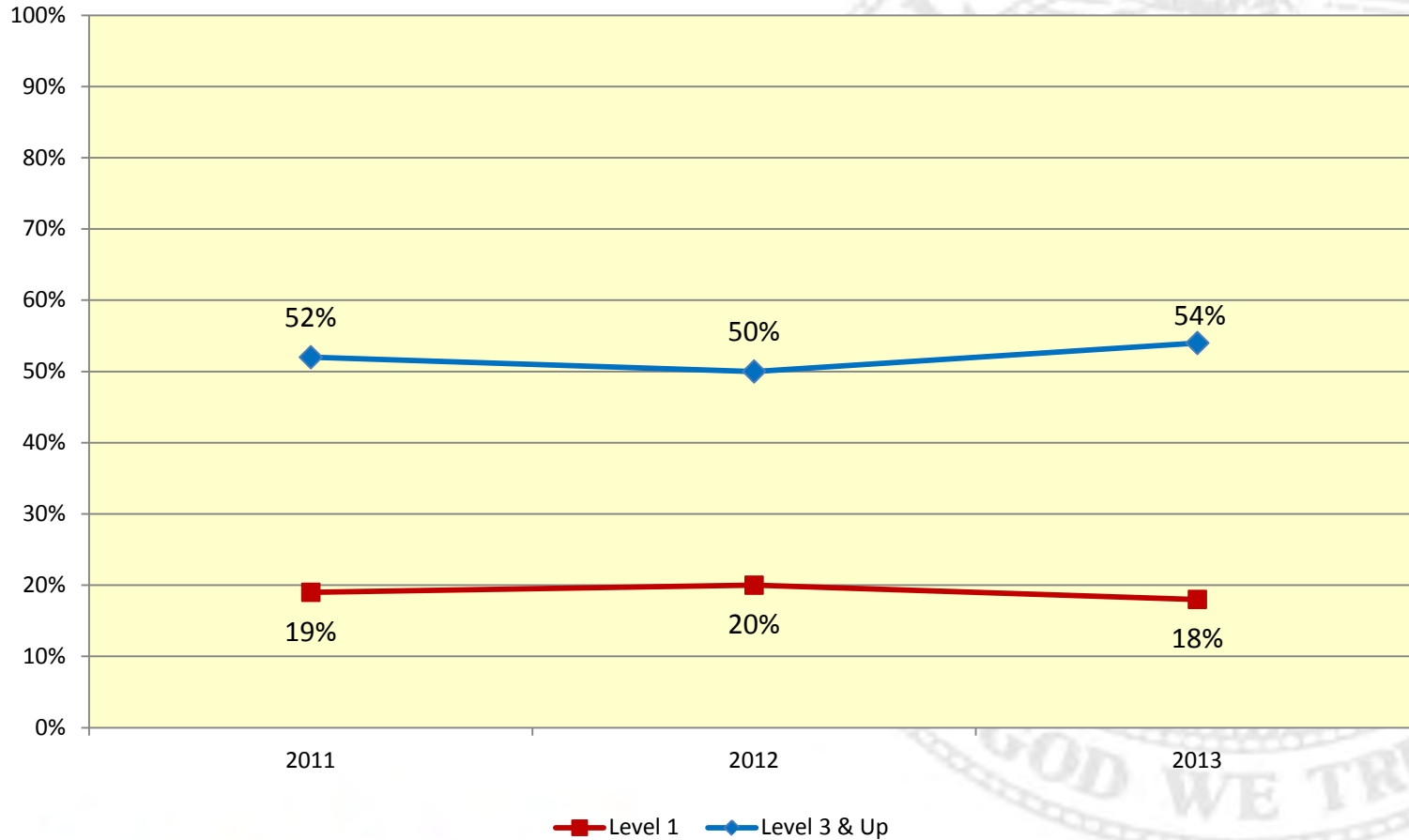
Note: <sup>1</sup> Accommodations were not permitted for this assessment.

# Grade 10 FCAT 2.0 Reading

## Historical Trend

- 2013 FL U.S. History EOC Assessment test-takers are mostly from grade 11

### Grade 10 FCAT 2.0 Reading





# Your Mission

- Separate these concepts
  - Achievement Levels
  - Accountability
- Focus on student achievement
  - Levels of success with the challenging content of the Next Generation Sunshine State Standards
- Set appropriately high standards for Florida's students

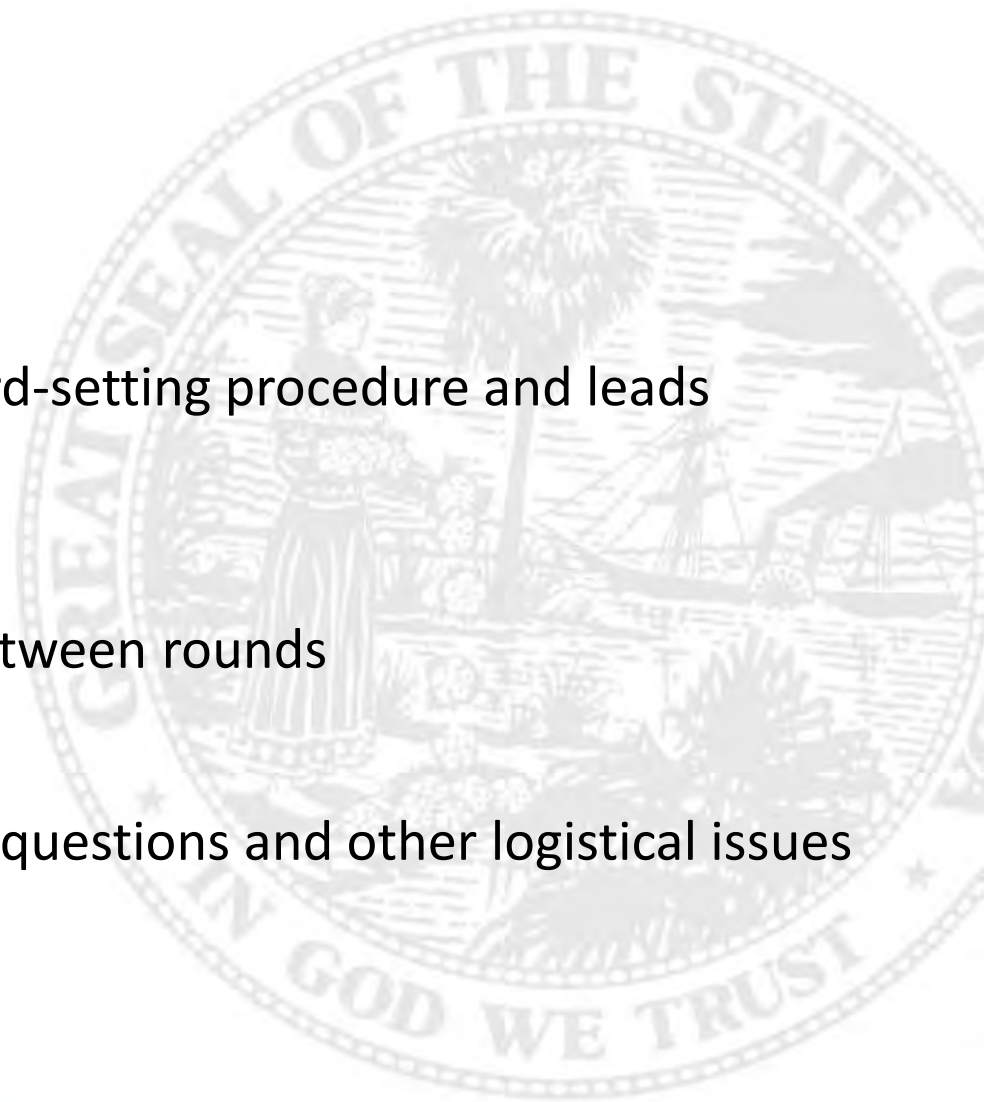


# Overview of the Process

- Educator judgments
- Superintendent/Business Community reactions
- Florida Department of Education reactions
- State Board of Education Rule advertised
- Public hearings to provide reactions
- Adoption by the State Board of Education

# Pearson Staff Roles

- Facilitator
  - Leads general session
  - Provides process oversight
  - Provides training on standard-setting procedure and leads breakout session activities
- Data Analyst
  - Computes feedback data between rounds
- Program Team Staff
  - Respond to reimbursement questions and other logistical issues

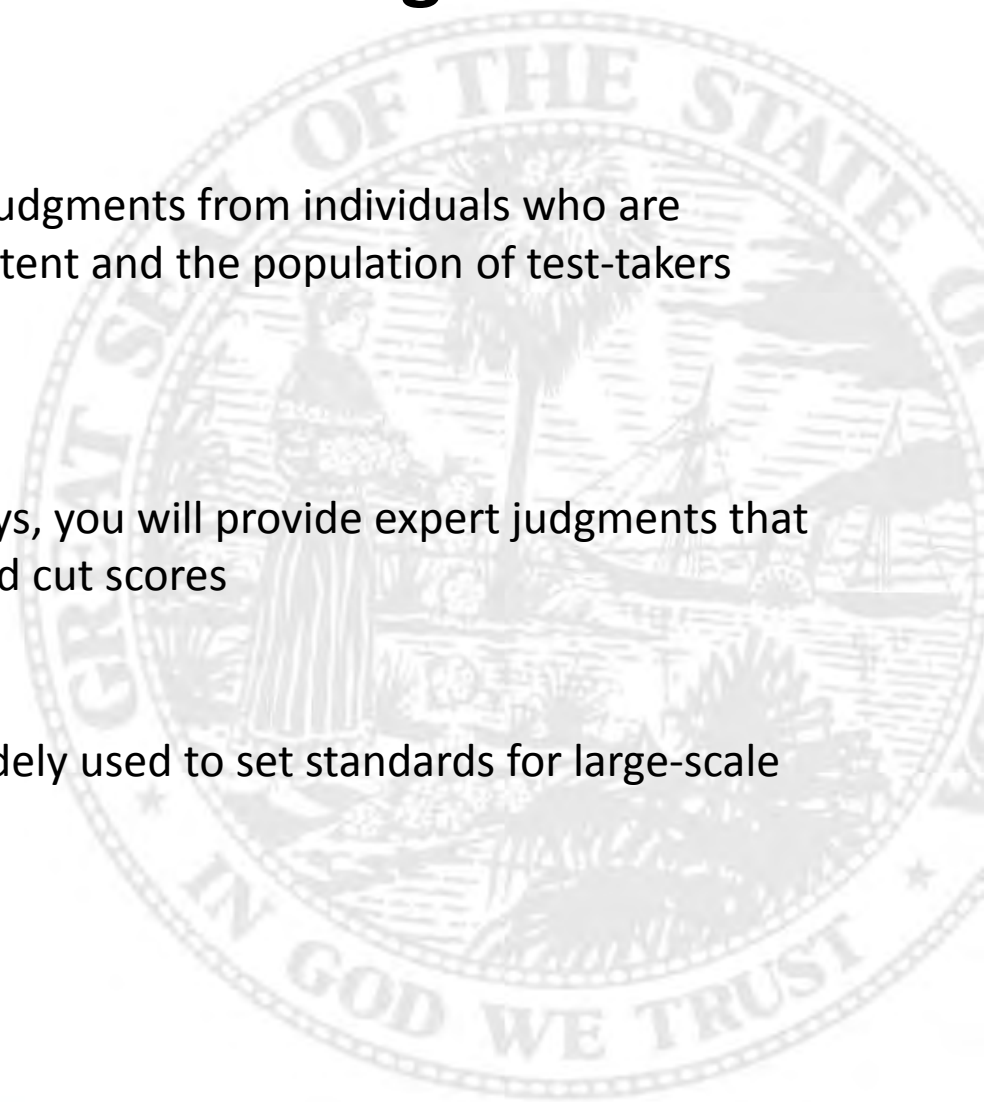


# Confidentiality

- DO NOT
  - Discuss the test items outside of this meeting
  - Remove any secure materials from the room on breaks or at the end of the day
  - Discuss judgments or cut scores (yours or others) with anyone outside of the meeting
  - Discuss secure materials with non-participants
  - Use cell phones in the meeting rooms (*Please turn your cell phone ringer off.*)
- “What happens in the meeting room stays in the meeting room.”
- General conversations about the process and days’ events are okay, but participants should avoid discussing details, particularly those involving items, cut scores, and any other confidential information
- Notes should be taken using provided materials only
- The only materials allowed on the table are standard-setting materials

# Purpose of the Meeting

- Why you are here
  - Standard setting relies on expert judgments from individuals who are knowledgeable about the test content and the population of test-takers
  - This is one step in a larger process
- What we will do
  - Over the next three-and-a-half days, you will provide expert judgments that will be used to form recommended cut scores
- How we will set standards
  - We will use a technique that is widely used to set standards for large-scale assessments





# When is Standard Setting Necessary?

- Standard setting becomes necessary whenever any of the following occur:
  - New test
  - Curriculum updates
  - Blueprint changes
  - Achievement Level Descriptions (ALDs) change
- Next Generation Sunshine State Standards – content standards

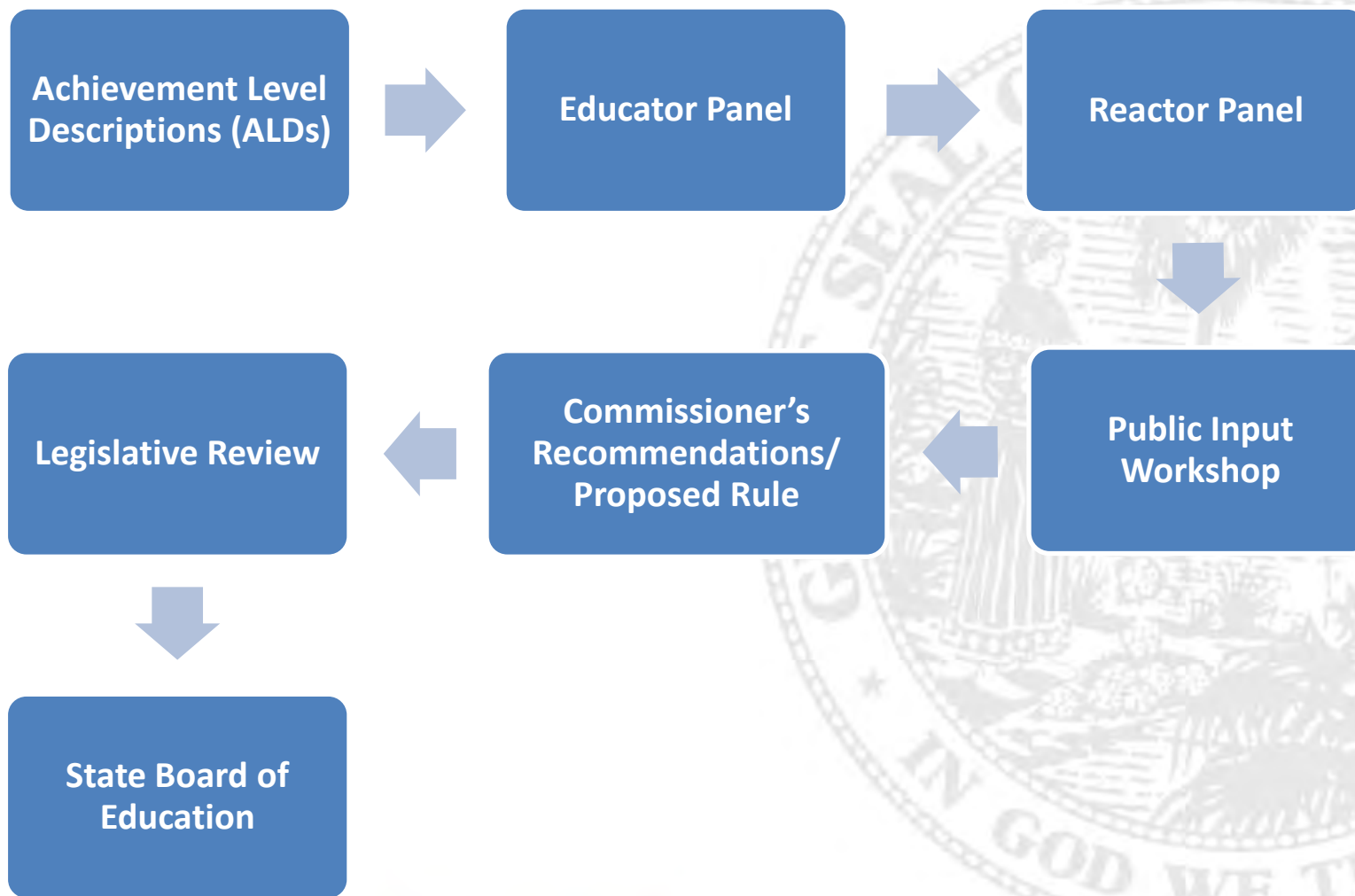


# Setting Standards is a Multi-Stage Process

Complex process with input solicited from several groups of stakeholders

- Completed: Content experts define Achievement Level Descriptions (ALDs). U.S. History EOC Assessment ALDs were drafted before August 2013.
- **August 13-16, 2013: Content experts rate the difficulty of items on the test relative to student expectations, which are aggregated to derive recommended cut scores.**
- August 22-23, 2013: Reactor panel reviews outcomes of standard setting meeting and provides feedback and recommendations for adopting the cut scores.
- Winter 2013-2014: State Board of Education reviews standard setting meeting outcomes and reactor panel feedback and makes final cut score decision.

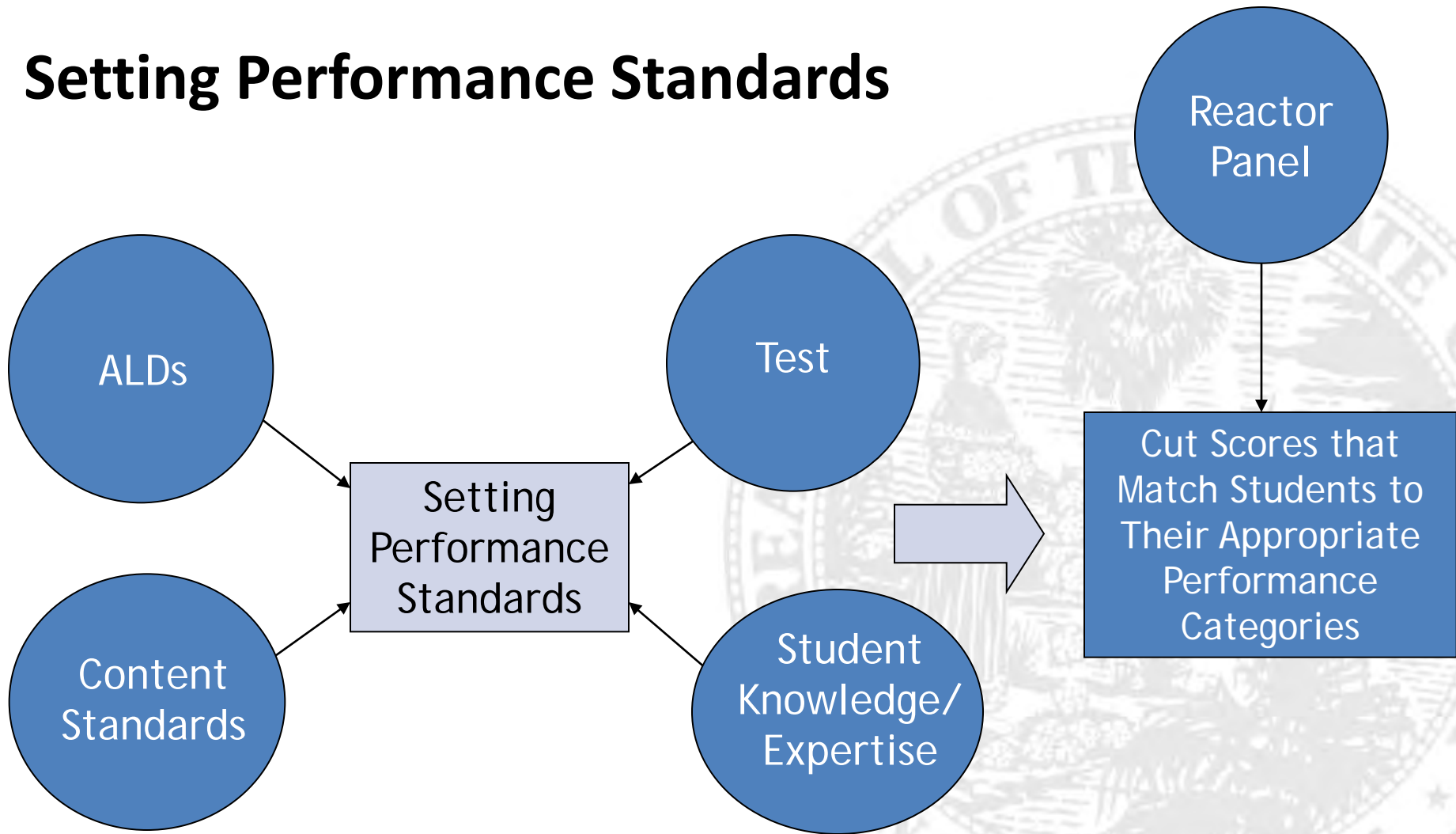
# Setting Standards is a Multi-Stage Process



# What is Standard Setting?

- A process of deriving levels of performance on educational or professional assessments, by which decisions or classifications of persons will be made (Cizek, 2006)
- Test scores can be used to group students into meaningful Achievement Levels
- Standard setting is the process whereby we “draw the lines” that separate the test scores into various Achievement Levels

# Setting Performance Standards



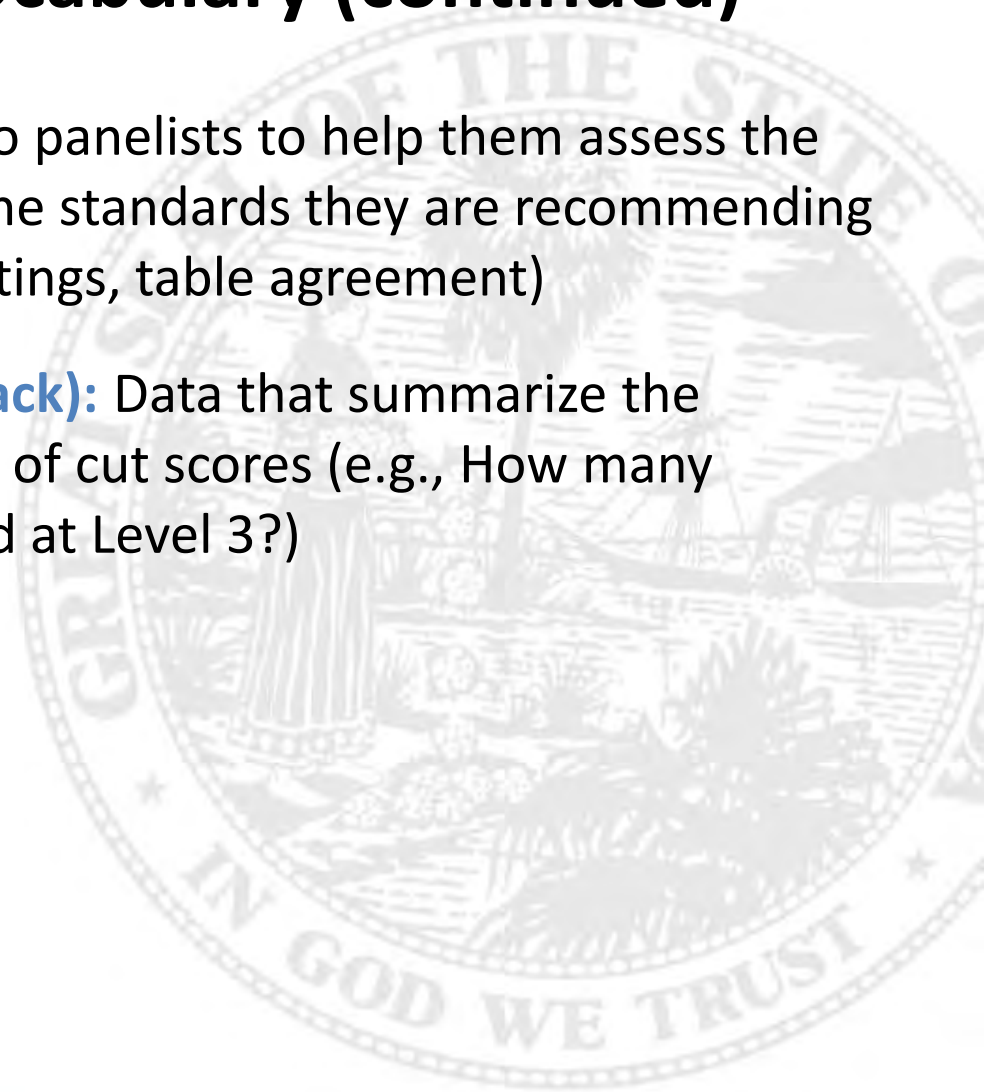


# Standard-Setting Vocabulary

- **Content Standards:** The content that students are expected to know
- **Achievement Levels:** Levels of student achievement based on observed scale scores
- **Achievement Level Descriptions (ALDs):** Descriptions of the competencies associated with each level of achievement
- **Cut Scores (Standards):** Scores on an assessment that separate one Achievement Level from another
- **Panelists (Judges/Raters):** Those who participate in the standard-setting process (stakeholders, educators, professionals – must understand the content assessed)

# Standard-Setting Vocabulary (continued)

- **Feedback Data:** Data provided to panelists to help them assess the validity and reasonableness of the standards they are recommending (e.g., median/mean cut score ratings, table agreement)
- **Impact Data (Normative Feedback):** Data that summarize the consequences of a proposed set of cut scores (e.g., How many students' scores will be classified at Level 3?)



# Standard-Setting Method

- One of a number of approaches available for setting standards
  - Judgmental procedure
- The standard-setting approach we will be following is one of the most widely-used methods for setting standards
- Judges consider characteristics of each item and expectations of test-takers to render item-level judgments that can be aggregated into overall cut scores for the test form
- Multiple rounds of judgments and delivery of information are designed to optimize decision making

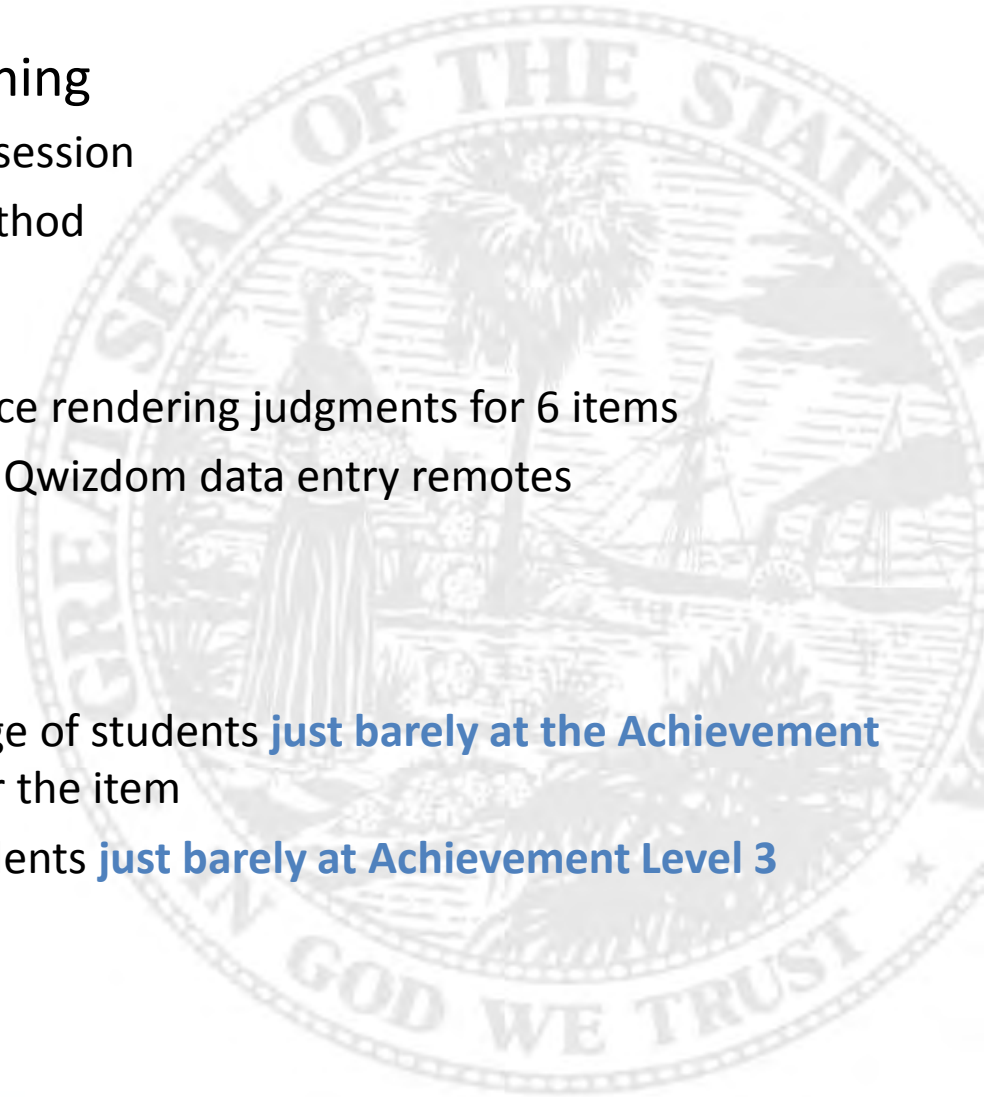
# Standard-Setting Procedure

- Review and discuss Achievement Level Descriptions (ALDs)
  - Developed prior to standard setting by Florida educators
  - Define expectations of students at each level
  - Distinguish students at adjacent levels
  - Panelists discuss characteristics that define students who are *just barely* at a given Achievement Level
  
- Panelists “take the test”
  - Panelists think about the test experience as if they were students
  - Consider the knowledge/skills required to answer each item correctly



# Standard-Setting Procedure (continued)

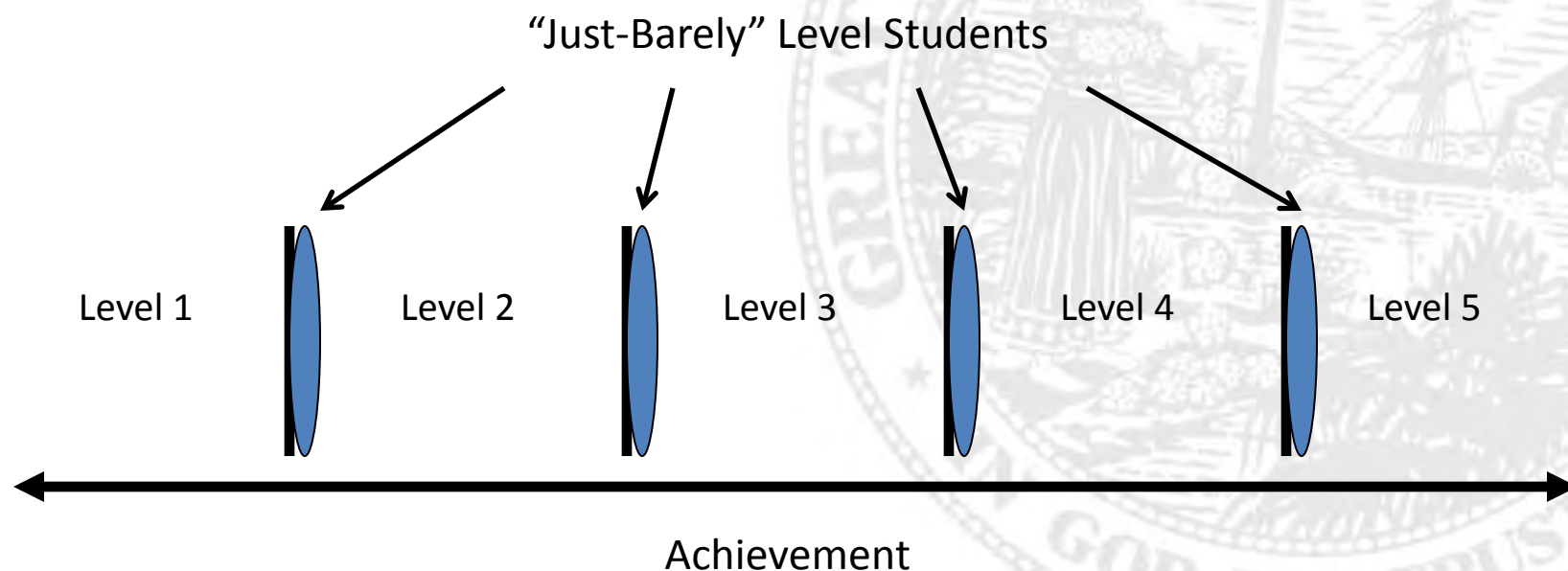
- Standard-setting procedure training
  - Provided later in standard-setting session
  - Details of the standard-setting method
- Practice Round
  - Opportunity for panelists to practice rendering judgments for 6 items
  - Practice entering judgments using Qwizdom data entry remotes
- Panelist Judgments
  - Panelists review each item
  - Estimate and record the percentage of students **just barely at the Achievement Level** who should correctly answer the item
  - Example: What percentage of students **just barely at Achievement Level 3** should answer item 1 correctly?





# The “Just-Barely” Test Taker

- Borderline in terms of Achievement Level
- Just barely meets criteria to be classified into the Achievement Level

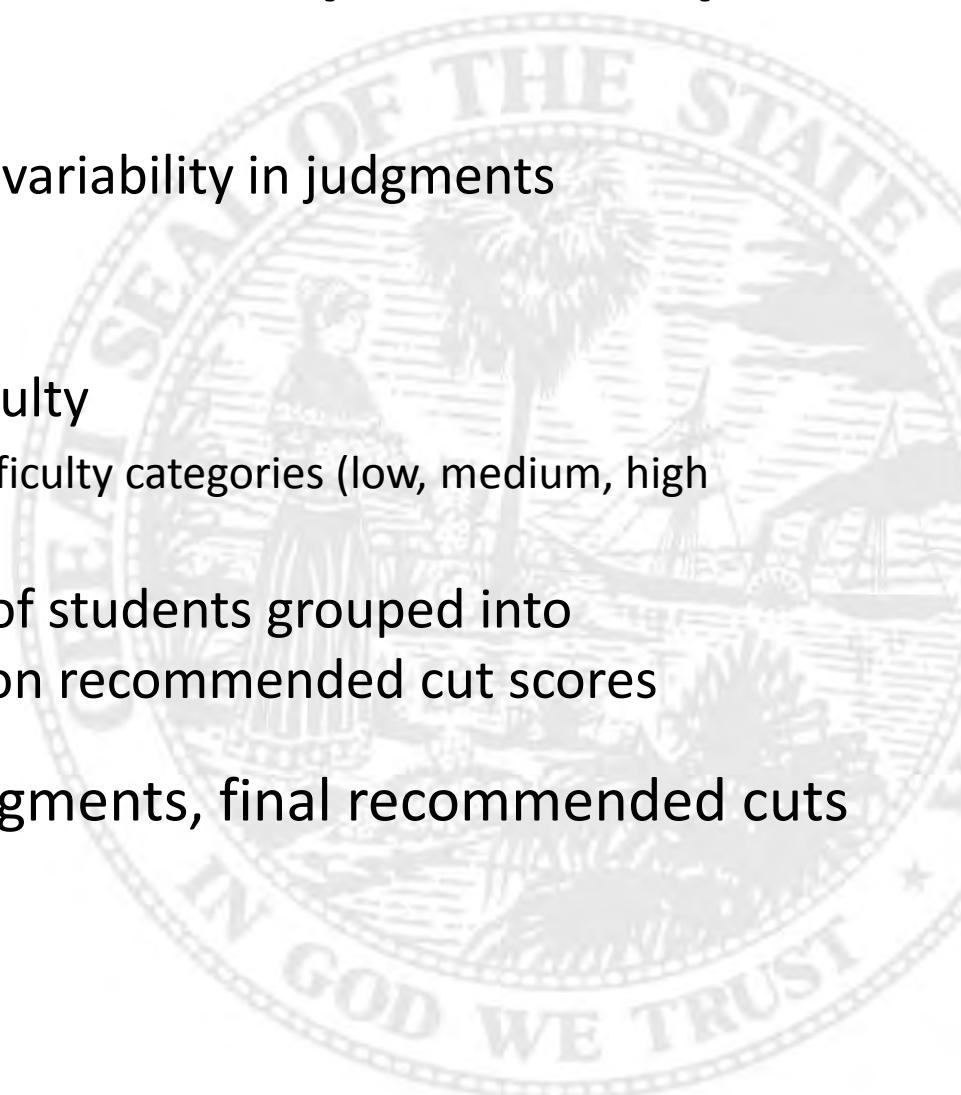


# Standard-Setting Procedure (continued)

- Recording judgments
  - Panelists provide separate judgments for each Achievement Level for each item (i.e., if there are four cut scores, the judge will rate each item four times)
  - Judgments recorded using electronic input device (clicker) and also recorded using paper and pencil
    - Clicker facilitates fast feedback data turnarounds
    - Paper/pencil recorded ratings help panelists track their judgments across Achievement Levels
  
- Multiple rounds of judgment
  - Panelists will render judgments across four rounds
  - Feedback provided between rounds
  - Feedback provided as a “reality check”
  - Discussion with peers between rounds of judgments

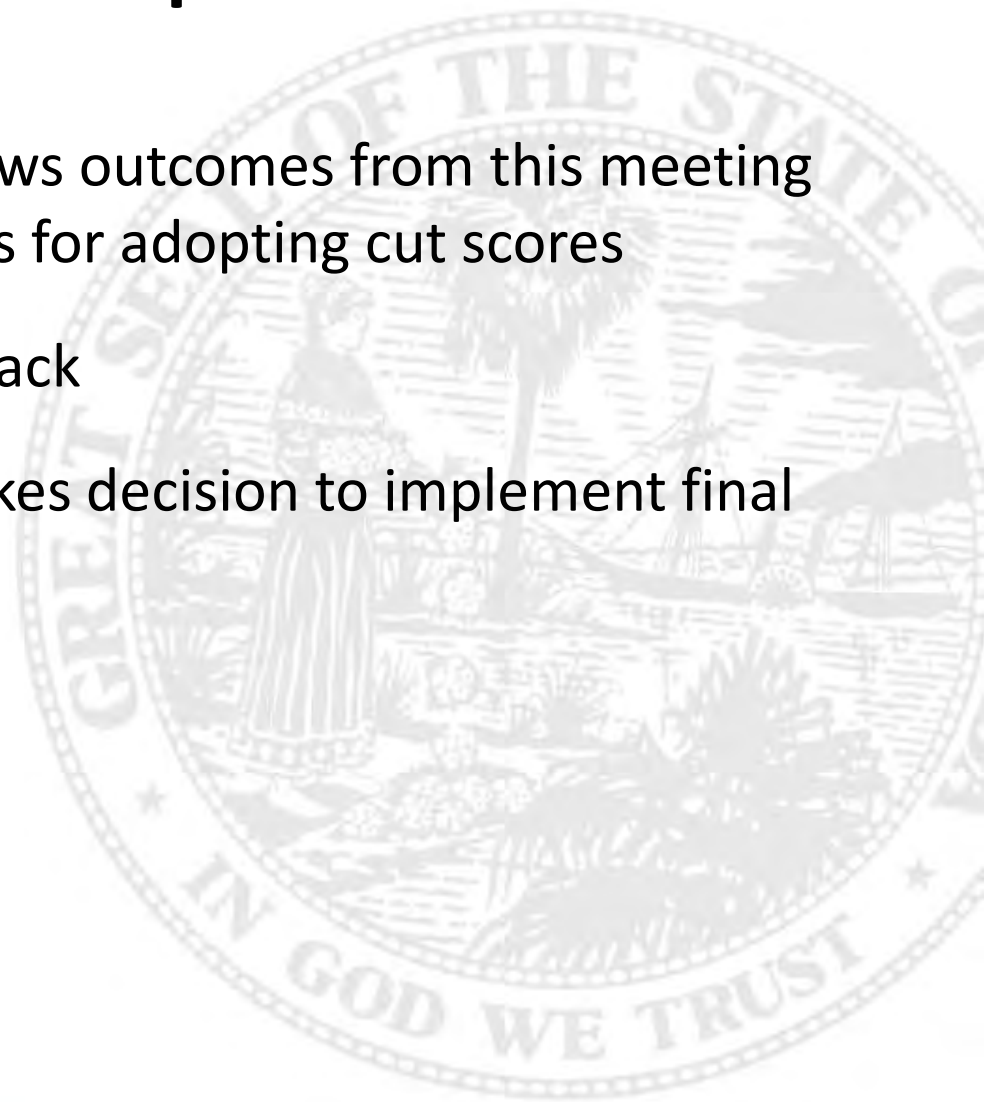
# Standard-Setting Procedure (continued)

- Feedback between rounds
  - Discuss items with greatest variability in judgments
    - Table discussion
    - Committee discussion
  - Review empirical item difficulty
    - Items classified into three difficulty categories (low, medium, high difficulty based on  $p$ -values)
  - Impact data – percentages of students grouped into Achievement Levels based on recommended cut scores
- Following Round 4 (final) judgments, final recommended cuts and impact data provided



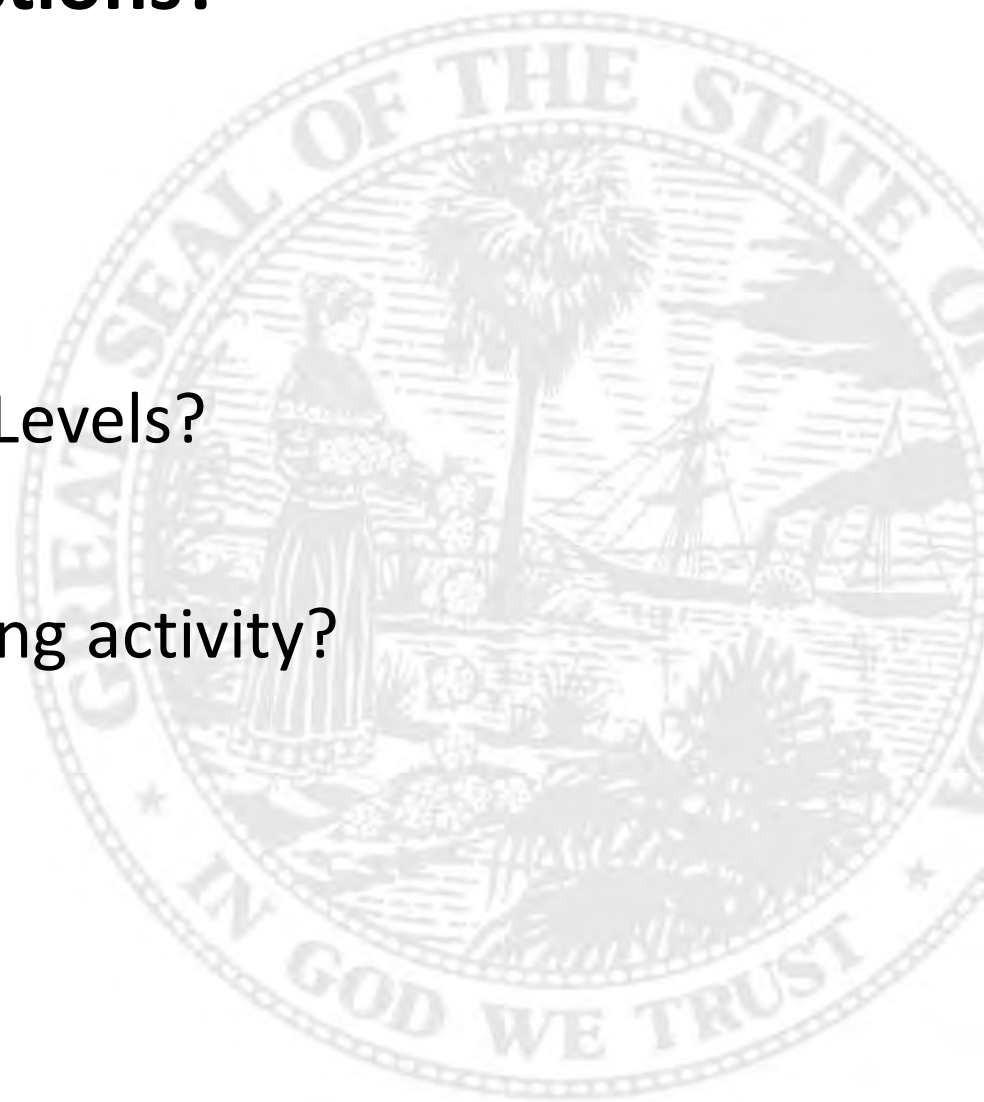
## Future Steps

- Separate Reactor Panel reviews outcomes from this meeting and makes recommendations for adopting cut scores
- Opportunity for public feedback
- State Board of Education makes decision to implement final cut scores



# Questions?

- About the tests?
- About the Achievement Levels?
- About the standard-setting activity?





# Time for a break!

Please return so we can begin promptly at 2:45 PM.

