



2015 Florida Standards Assessments Standard Setting Educator Panel

General Session

Educator Panel

August 31, 2015



Presenters

- Vince Verges, Assistant Deputy Commissioner, Division of Accountability, Research, and Measurement, Florida Department of Education
- Victoria Ash, Chief, Bureau of K–12 Student Assessment, Florida Department of Education
- Gary W. Phillips, Vice President and Institute Fellow, American Institutes for Research



Overview—What Are You Doing Here?

- To thoroughly review the content requirements of the Florida Standards Assessments (FSA)
- To help the State of Florida establish achievement level standards for these assessments



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”)
 - Sunshine State Standards (FCAT)
 - Next Generation Sunshine State Standards (FCAT 2.0)
 - **Florida Standards (FSA)**
- Achievement: Describe how much content knowledge a student is required to demonstrate
 - Achievement Level Standards
 - Graduation Requirement (Algebra 1 and Grade 10 ELA)
- Accountability Standards
 - School Grading Criteria
 - Annual Measurable Objectives



FSA Grade-Level/Subject Tests

- Grades 3–10 English Language Arts
- Grades 3–8 Mathematics
- Algebra 1, Algebra 2, Geometry End-of-Course (EOC) Assessments



Administration Information

Grade/Subject Assessment	Session Length	Number of Sessions
Grades 4–10 ELA Writing	120 Minutes	1
Grades 3–5 ELA Reading	80 Minutes	2
Grades 6–8 ELA Reading	85 Minutes	2
Grades 9–10 ELA Reading	90 Minutes	2
Grades 3–5 Mathematics	80 Minutes	2
Grades 6–8 Mathematics	60 Minutes	3
Algebra 1, Geometry, Algebra 2 EOCs	90 minutes	2



Administration Information

- Except for ELA Writing (one session only), all assessments are administered over two days
- Grades 4–7 ELA Writing and Grades 3–4 ELA Reading and Mathematics are administered as paper-based tests; all other assessments are computer-based.
- Paper-based accommodations are offered for eligible students.
- Students taking paper-based and computer-based tests had opportunities to access practice tests to become familiar with item types, functionality, and test mode or platform.



Your Responsibilities

- Focus on student achievement
 - Levels of success with the challenging content of the Florida Standards
- Set appropriately high standards for Florida's students



Assessment Contractor Staff Roles

- American Institutes for Research (AIR) Facilitator
 - Leads general session
 - Provides process oversight
 - Provides training on standard-setting procedure and leads breakout session activities
 - Computes feedback data between rounds
- Data Recognition Corporation (DRC) 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 acceptable, 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 four 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

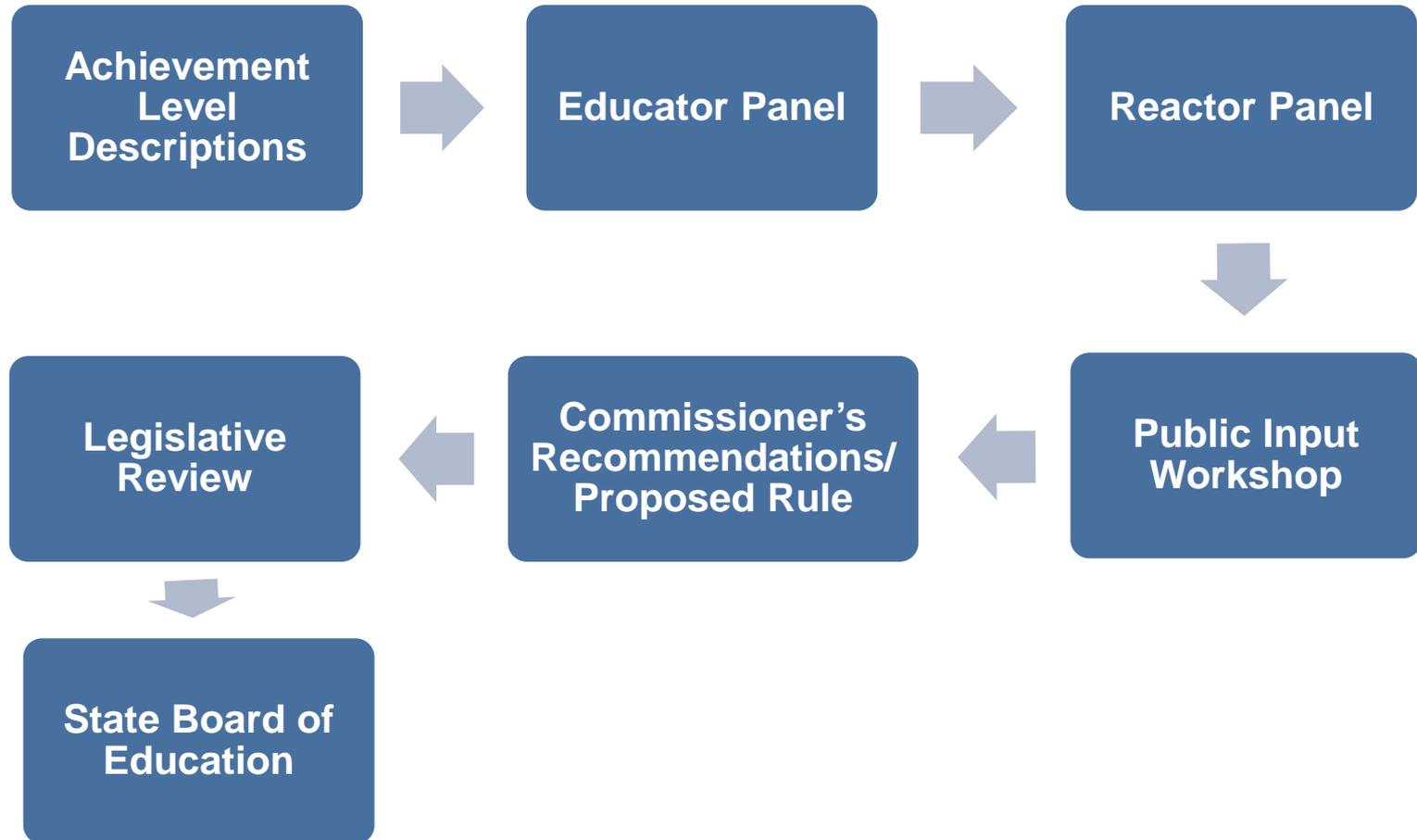


When Is Standard Setting Necessary?

- Standard setting becomes necessary whenever any of the following occur:
 - New test
 - Curriculum updates
 - Blueprint changes
 - ALDs change
- FSA: new assessments for adopted content standards



Stages in the Standard Setting Process





Important Dates in Multi-Stage Process

- Achievement Level Description (ALD) Panel:
April 28–May 1, 2015
 - Hotel Duval, Tallahassee, Florida
- Educator Panel: August 31–September 4, 2015
 - Grand Cypress, Orlando, Florida
- Reactor Panel: September 10–11, 2015
 - Grand Cypress, Orlando, Florida
- Fall/Winter 2015: Legislative review and input period completed.
- Fall/Winter 2015: State Board of Education reviews standard-setting meeting outcomes and reactor panel feedback and makes final cut score decision.



Achievement Level Description Panel

- April 28–May 1, 2015
- Four-day workshop
- Forty-two panelists



Educator Panel

- August 31–September 4, 2015
- Five-day standard-setting workshop
- Four rounds of standard setting
- Approximately 300 panelists
- Seventeen rooms setting standards concurrently



Reactor Panel

- September 10–11, 2015
- Two-day meeting composed of
 - community/education organization leaders
 - state university leaders
 - business leaders
 - school board members
 - superintendents
- Two rounds of review
- 16 panelists

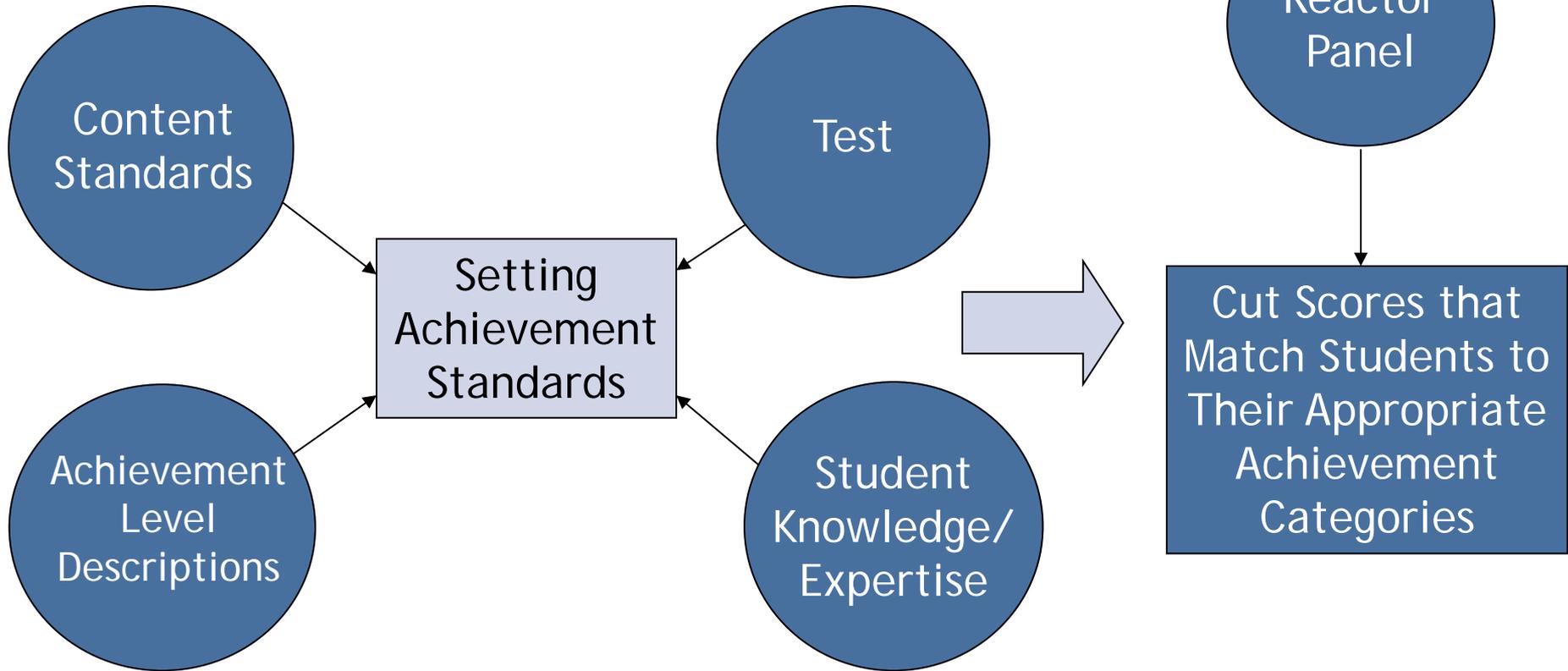


What Is Standard Setting?

- 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 Achievement Standards





Overall Structure of the Educator Panel

Panel	Panelists	Subject	Grade/EOC	AIR Facilitator	AIR Facilitator Assistant
1	18	Mathematics	3	Alysa Kartee	Tiffany Abu-Shaikha
2	19	Mathematics	4	Jim McCann	Lisa Schaaf
3	21	Mathematics	5	Paul Maxon	Daniel Freedberg
4	18	Mathematics	6	Erica Ajder	Eileen Heneghan
5	20	Mathematics	7	Maureen Font	Nate Thompson
6	16	Mathematics	8	Jennifer Rubel	Bernard Farley
7	19	Algebra 1	EOC	Kari Stellpflug	Christina Estes
8	16	Geometry	EOC	Sam Thomas	Marie Kristine-Tardif
9	16	Algebra 2	EOC	Chris Paskoff	Susan Sherwood
10	17	ELA	3	Allison Stingley	Stephanie Ryan
11	19	ELA	4	John Neral	Jacob Wilkes
12	18	ELA	5	Sean Redmond	Kevin Clayton
13	21	ELA	6	Brett Craycraft	Sarah Abdelnaby
14	15	ELA	7	Diana Reed	Terra Winsett
15	16	ELA	8	Natalie Rebentisch	Amber Benlian
16	15	ELA	9	Kelly Quinney	Anthony Kazanjian
17	18	ELA	10	Katina Marshall	Brian Kline
				June Zack	Mathematics Lead
				Kevin Dwyer	ELA Lead



Main Activities of the Educator Panel

- Table leader training
- Large group orientation
- Grade/subject specific training
- Panelists will
 - take the online test in subjects/grades that are online
 - review the content standards
 - review achievement level descriptions
 - create just barely summary ALDs
 - review the ordered item booklet
- Recommend four achievement standards in four rounds
 - Level 2, Level 3, Level 4, and Level 5
- Workshop evaluation



What Will Be Online for the Educator Panel?

1. Tests for grades/subjects administered online
2. Ordered item booklet
3. Bookmark placements (four rounds)
4. Feedback results
5. Impact data
6. Benchmark data



Computer-Based and Paper-Based Tests by Grade and Subject

Grade/Subject Test	Paper-Based	Computer-Based
Grade 3 ELA	Reading (no Writing administered)	
Grade 4 ELA	Reading and Writing	
Grades 5-7 ELA	Writing	Reading
Grades 8-10 ELA		Reading and Writing
Grades 3-4 Mathematics	✓	
Grades 5-8 Mathematics		✓
EOCs (Algebra 1, Geometry, Algebra 2)		✓



Vertical Scales and Within-Grade Scales

- Vertical scales
 - English language arts scale, grades 3–10
 - Mathematics scale, grades 3–8
- Within-grade scales
 - Algebra 1
 - Algebra 2
 - Geometry



What Method are we using for Standard Setting?

- Bookmark Method*
 - Research-based procedure
 - Used in many state assessment programs
 - Proven to be technically sound in litigation

*Mitzel, H. C., Lewis, D. M., Patz, R. J. & Green, D. R. (2001). "The Bookmark procedure: Psychological perspectives." In G. Cizek (ed.), *Setting performance standards: Concepts, methods and perspectives*. Mahwah, NJ: Erlbaum.

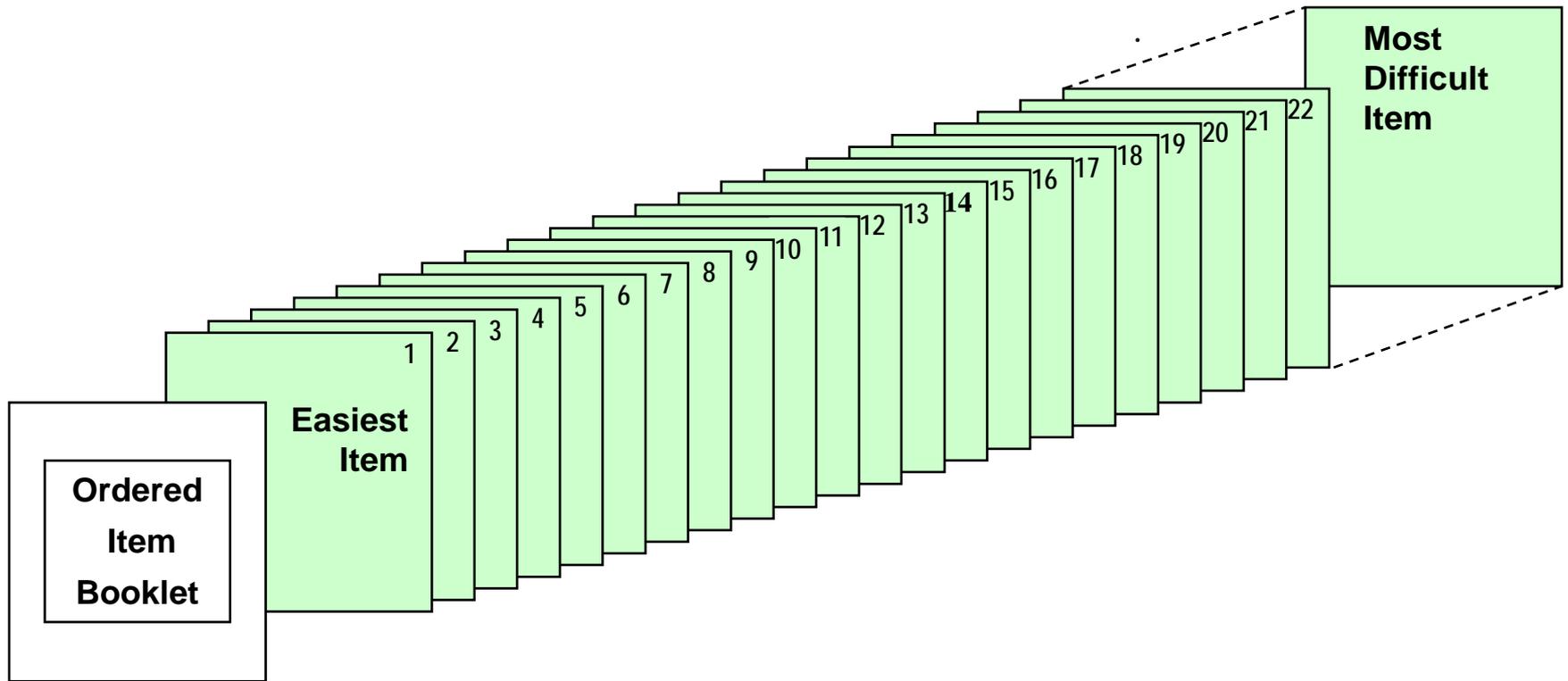


Bookmark Method using Ordered Item Booklet (OIB)

- Items are ordered by difficulty
- Each page is a score point on an item
- Some constructed-response items appear multiple times in OIB (once for each score point)
- The number of pages in the book is equal to the number of points in the ordered item booklet (not the number of items)

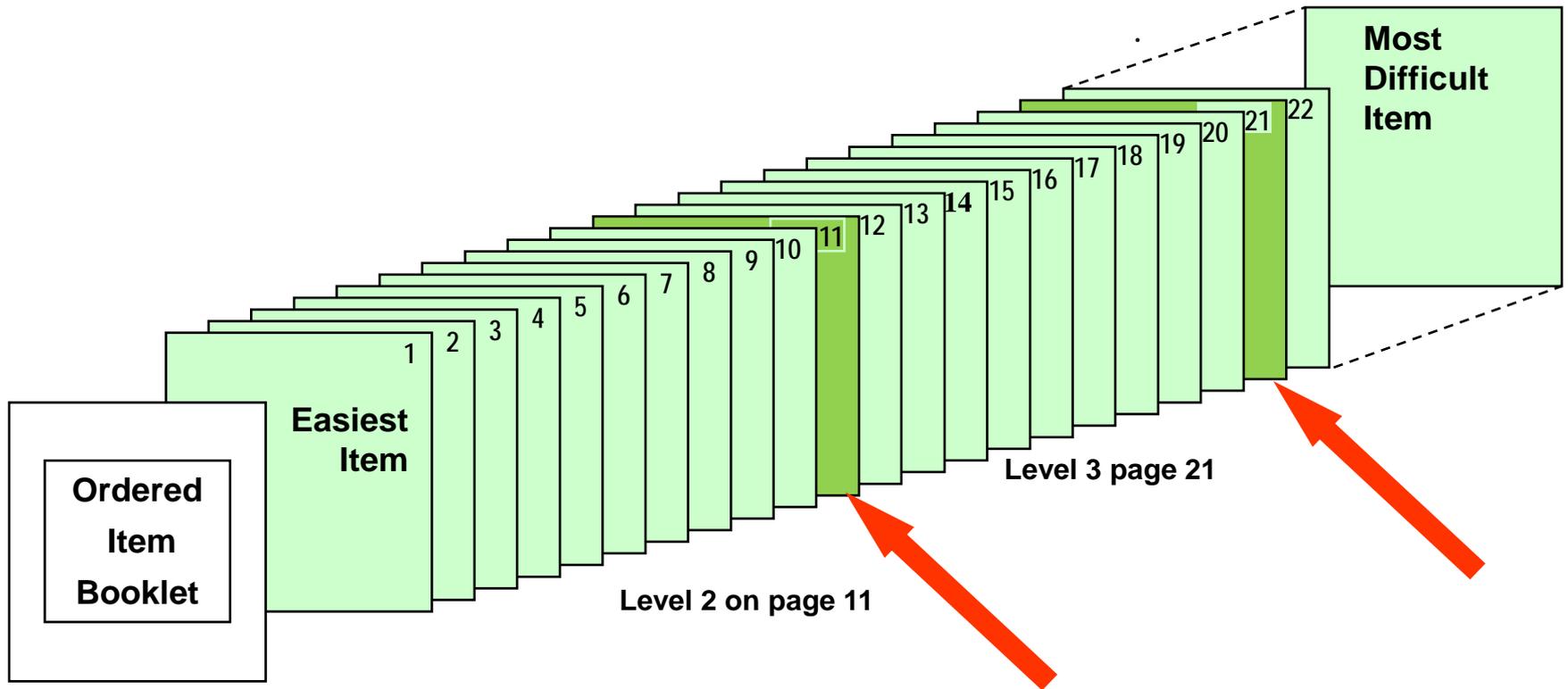


Ordered Item Booklet (OIB)





Bookmarking Pages in the Ordered Item Booklet





Bookmarking Pages in the Ordered Item Booklet

- The terms “bookmarking pages” and “ordered item booklet” historically come from a pencil/paper testing environment
- For the FSA, the ordered item booklet will be online so the pages you select will be from a drop-down menu



Day 1 Summary

- Overview of standard-setting process
- Take the test
- Review achievement level descriptions
- Create just barely achievement level descriptions



Day 1

Overview: What Are Content Standards?

- Florida Standards
- Specify what students know and can do
- Can be found at <http://www.flstandards.org/>



Day 1

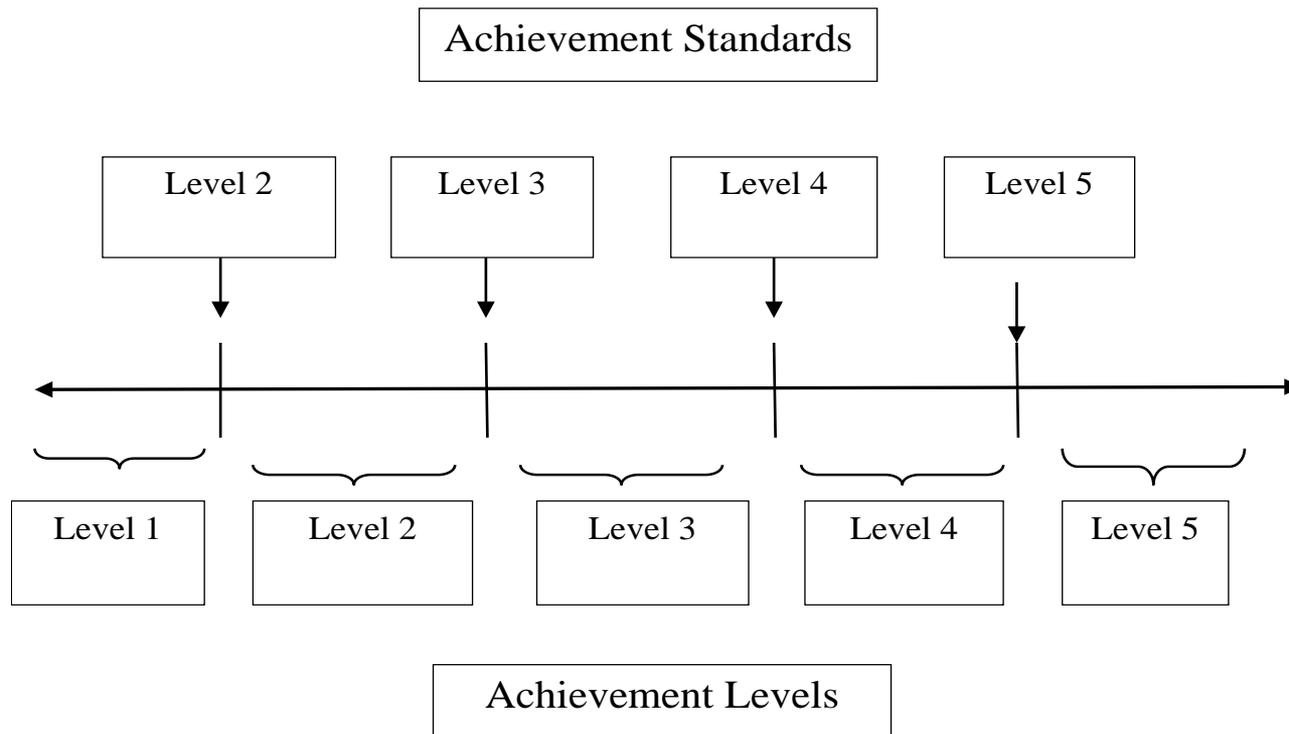
Overview: What Are Achievement Standards?

- Specify how many of the content standards students must know and be able to do in order to meet each achievement level
 - Four achievement standards (cut scores)
 - Five achievement levels



Day 1

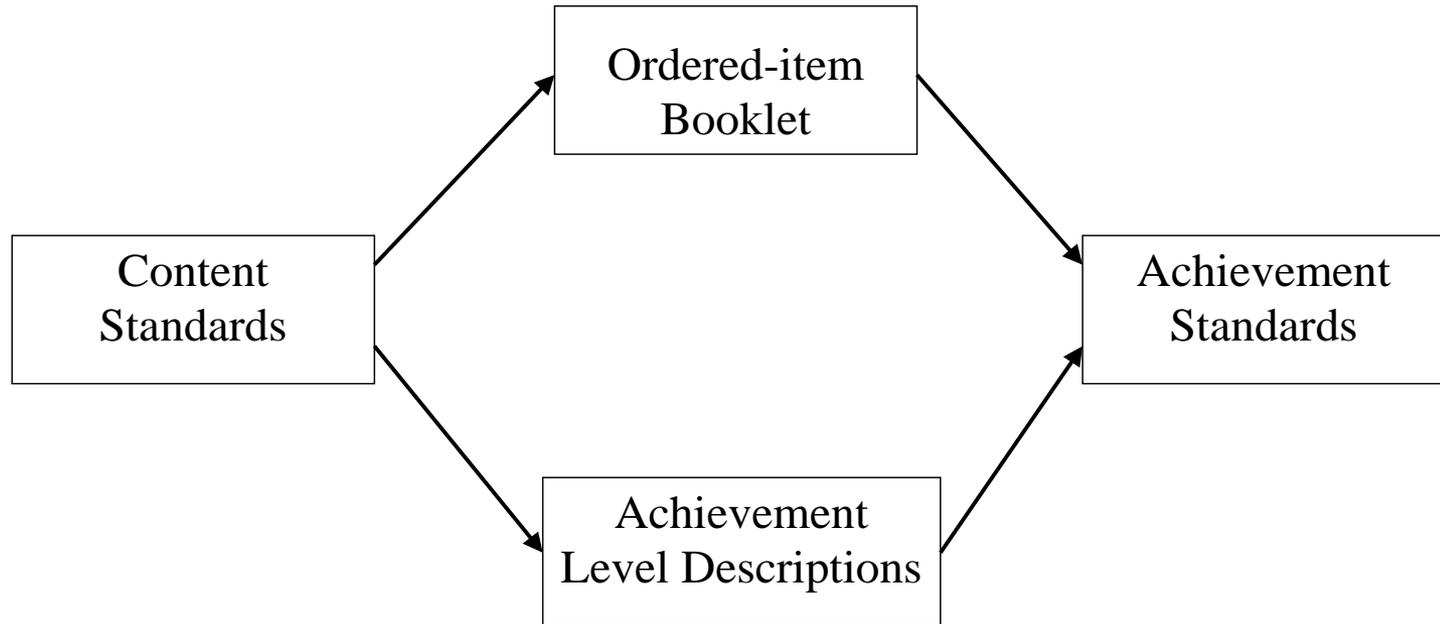
Overview: What Are Achievement Standards?





Day 1

Overview: From Content Standards to Achievement Standards





Day 1

Take the Test

- Items administered in spring 2015
- For computer-based tests, interface is identical to the online test environment that students experienced
- This is an opportunity to interact with the items



Day 1

Achievement Level Descriptions (ALDs)

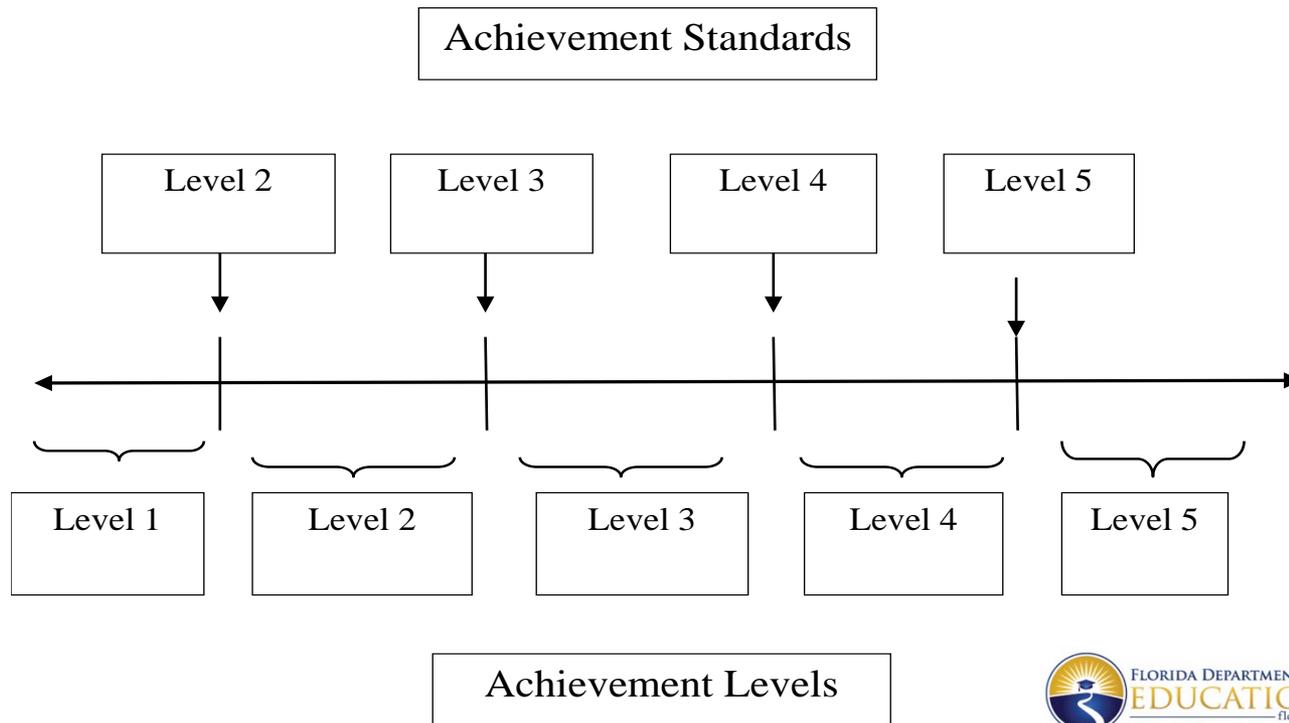
- Specify what students in each achievement level are expected to know and be able to do
- ALDs are the link between content and achievement standards
- Use the ALDs to develop a mental representation of students at each achievement level
- Place the bookmark at the point where students scoring at and above that level can be accurately described by the ALD



Day 1

Just Barely Achievement Level Description

- When considering each achievement level, we are interested in those students who just barely reach the standard
- Not typical of students in achievement level. Although just barely, they do reach the standard.

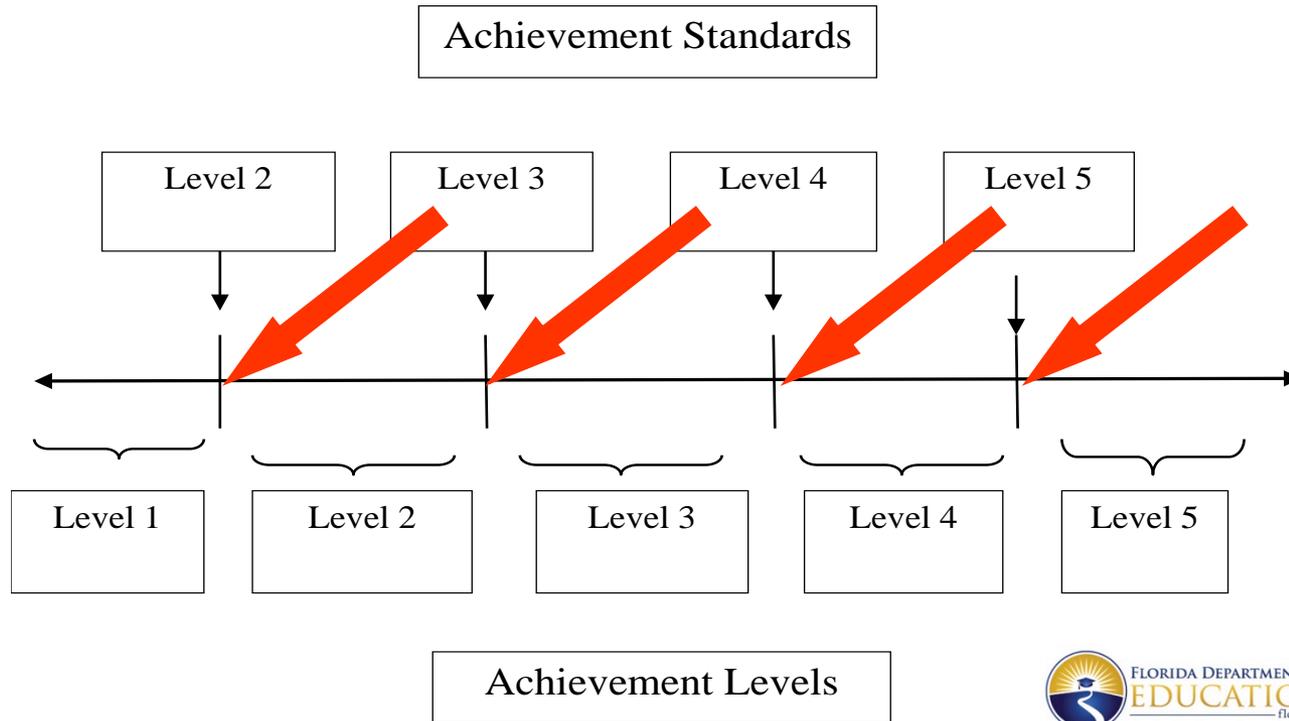




Day 1

Just Barely Achievement Level Description

- When considering each achievement level, we are interested in those students who just barely reach the standard
- Not typical of students in achievement level. Although just barely, they do reach the standard.





Day 1

Achievement Level Descriptions (ALDs)

- Policy ALDs
- Range ALDs
- Just barely ALDs
- Reporting ALDs

NOTE: Each of the 17 rooms will have its own ALDs.



Day 2 Summary

- Review the Bookmark method of standard setting
- Review the ordered item booklet
- Practice test
- Complete readiness form
- Round 1 recommendations for
 - Level 3
 - Level 2
 - Level 4
 - Level 5



Day 2

Studying the Ordered Item Booklet

- Consider each item and answer two questions:
 1. What do students need to know and be able to do to respond successfully to this item?
 2. Why is this item more difficult than the previous items?



Day 2

What If an Item Seems Out of Order?

- Item order is based on student performance
- Items may seem out of order because they are ordered by difficulty, not by content or cognitive process
- The sequence of items in the OIB will not match sequence of instruction taught throughout the school year
- The ordering of items in the OIB will not match the ordering of items on the test



Day 2

What If an Item Seems Wrong or Unfair?

- All items have been vetted and approved through content and bias and sensitivity reviews
- This is not another item review meeting
- If you believe something is wrong with an item, tell the AIR facilitator or an assistant facilitator; then skip the item as you review the rest of the OIB



Day 2

Remember Standard Setting Is Aspirational

- Standard setting is all about what students **should** know and be able to do, not about what they actually know and are able to do
- **Do not set standards for your classroom.** You are setting standards for all students across the state.



Day 2

Mechanics of Bookmark Procedure

- Initial judgment based solely on OIB (round 1)
- Articulation (introduced after round 1)
- Impact data (introduced after 2)
- Benchmarking (introduced after round 3)



Day 2

Getting Ready to Recommend Standards

- Practice using the OIB
- Filling out your readiness form



Day 2

Accessing the Ordered Item Booklet

- Open the Chrome browser
- Sign in with your user name and password

AIR Assessment[®]

Setting Standards: Log In

User Name:

Password:

Remember me next time.

Please log in



Day 2

- Submitted recommendations for round 1



Day 3 Summary

- Review feedback from round 1
- Review articulation information
- Round 2
- Feedback from round 2



Feedback

More about this item Notes Marks Impact **Feedback** Moderation Prior Feedback

Summary of tentative standards

Exhibit 1: Pages corresponding to room and table medians

	Level 2 - Below Satisfactory	Level 3 - Satisfactory	Level 4 - Above Satisfactory	Level 5 - Mastery
Room	14	24	45	55
Table 1	14	24	45	55
You	14	33	45	56

Footnote goes here

Exhibit 2: Bookmarks placed by panelists

Table	First Name	Last Name	Level 2 - Below Satisfactory	Level 3 - Satisfactory	Level 4 - Above Satisfactory	Level 5 - Mastery
1	Silvia	Aday	14	33	45	56
1	James	Chinn	13	24	35	46
1	Jonetta	Dawson	15	24	45	55



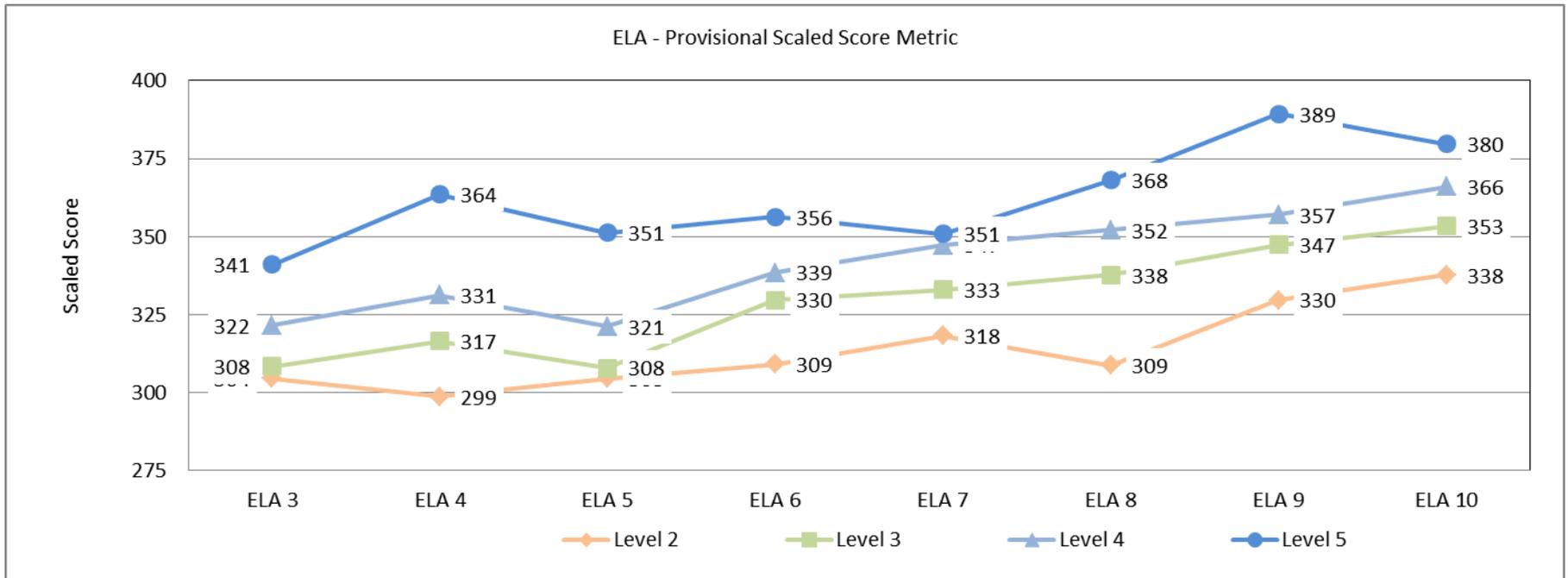
Day 3

Articulation

- Even though each room is recommending standards for one grade and one subject, the entire workshop is recommending a system of 17 sets of achievement standards.
- Achievement standards for a statewide system must be coherent across grades and subjects:
 - Vertically articulated across grades with no anomalous peaks and valleys
 - Orderly across subjects with no dramatic differences in expectation

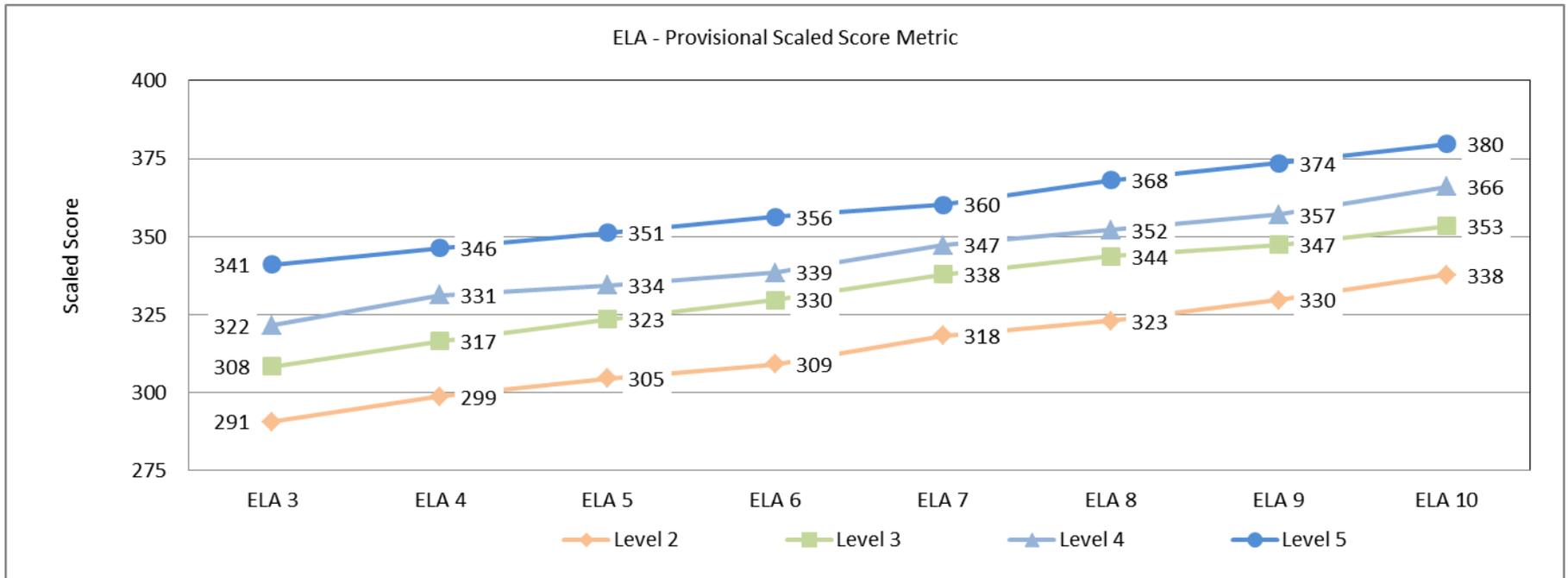


Day 3 - Illustration of Disarticulated Standards





Day 3 - Illustration of Articulated Standards





Day 3

Articulation

- AIR will estimate page numbers that would represent articulated standards.
- The articulated standards would be communicated to the panelists at the beginning of rounds 2, 3, and 4.
- With articulated cut scores in hand, the judgment task for panelists is modified so that panelists are now asked to judge whether it makes sense from a content point of view to place their bookmark on near the OIB page associated with each articulated cut score.



Day 3

- Submit recommendations for round 2



Day 4 Summary

- Discuss concept of impact data
- Round 3
- Feedback from round 3
- Large group discussion:
 - Status check on all standards
 - Discuss external benchmark data
- Individual room discussion



Day 4

Discuss Concept of Impact Data

- Shows the percent of students who would reach any standard that you select.
- Introduced after round 2 after the panelists have made judgments across two rounds based solely on content considerations.
- Impact data are used as context to inform the panelists' recommendations but should not determine their recommendations.
- In the end the panelists' recommendations should have a content justification.



Day 4

How Do We Display the Impact Information?

- As the panelists scroll through the online OIB they will be shown the impact percentages associated with each page.
- A graph shows the percent of students that would score at and above the achievement standard.
 - The entire room selected
 - Your table selected
 - You selected



Impact Data

More about this item Notes Marks **Impact** Feedback Moderation Prior Feedback

Some facts about the difficulty of this item (28)

Feedback	Value
Overall percent of Florida students that perform at or above this level:	43
Percentage of Male Florida students that perform at or above this level:	42
Percentage of Female Florida students that perform at or above this level:	45
Percentage of White Florida students that perform at or above this level:	53
Percentage of African American Florida students that perform at or above this level:	26
Percentage of Hispanic Florida students that perform at or above this level:	40
Percentage of American Indian Florida students that perform at or above this level:	42
Percentage of Asian Florida students that perform at or above this level:	73
Percentage of Pacific Islander Florida students that perform at or above this level:	58
Percentage of Multiracial Florida students that perform at or above this level:	48



Impact Data

More about this item Notes Marks Impact **Feedback** Moderation Prior Feedback

Summary of tentative standards

Exhibit 1: Pages corresponding to room and table medians

	Level 2 - Below Satisfactory	Level 3 - Satisfactory	Level 4 - Above Satisfactory	Level 5 - Mastery
Room	14	24	45	55
Table 1	14	24	45	55
You	14	33	45	56

Footnote goes here

Exhibit 2: Bookmarks placed by panelists

Table	First Name	Last Name	Level 2 - Below Satisfactory	Level 3 - Satisfactory	Level 4 - Above Satisfactory	Level 5 - Mastery
1	Silvia	Aday	14	33	45	56
1	James	Chinn	13	24	35	46
1	Jonetta	Dawson	15	24	45	55

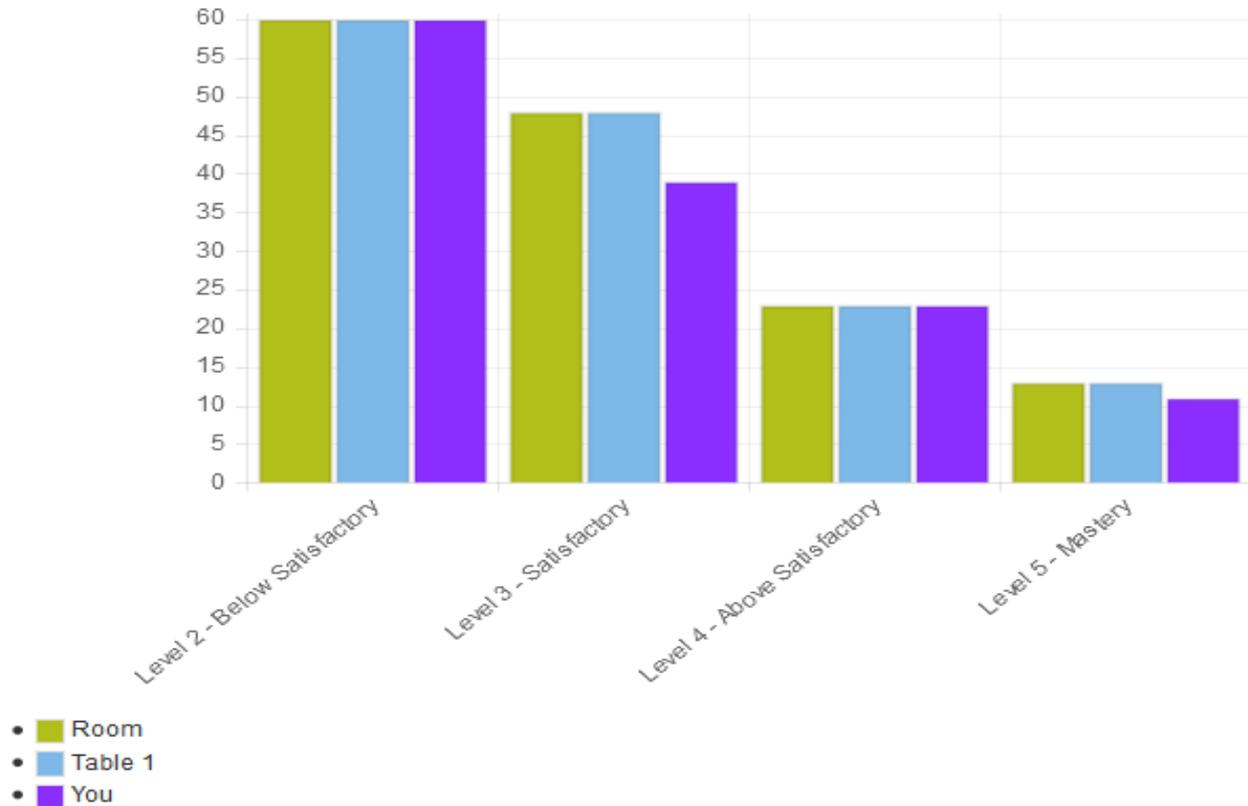
Exhibit 3: Impact corresponding to table and room medians: Overall percent of Florida students that perform at or above this level:

	Level 2 - Below Satisfactory	Level 3 - Satisfactory	Level 4 - Above Satisfactory	Level 5 - Mastery
Room	60	48	23	13
Table 1	60	48	23	13
You	60	39	23	11



Impact Data

Chart 3: Impact corresponding to table and room medians: Overall percent of Florida students that perform at or above this level:





Day 5 Summary

- Discuss concept of benchmark data
- Round 4
- Feedback from round 4
- Room discussion
- Adjourn



Day 5

Discuss Concept of Benchmark Data

- Benchmarking provides the panelists with external references so they can see how their recommendations compare with national and international standards.
- Benchmarking helps determine whether FSA achievement standards are nationally and internationally competitive
 - National Assessment of Educational Progress (NAEP)
 - Trends in International Mathematics and Science Study (TIMSS)
 - Program for International Student Assessment (PISA)



Benchmark Data

[More about this item](#)
[Notes](#)
[Marks](#)
[Impact](#)
[Feedback](#)
[Moderation](#)
[Prior Feedback](#)

Some facts about the difficulty of this item (42)

Feedback	Value
Overall percent of Florida students that perform at or above this level:	43
NAEP Florida: This level of achievement is classified as	On track to be at the Proficient Level on NAEP in Mathematics
TIMSS National: This level of achievement is classified as	On track to be at the High International Benchmark on TIMSS in Mathematics



Day 5

NAEP Benchmarks

- Florida NAEP results from the 2013 assessment
- Administered by the National Center for Education Statistics (NCES)
- Two-year assessment cycle
- Grades 4, 8, and 12
 - Grade 12 NAEP benchmark does not apply to the FSA because FSA ELA is grades 3–10 and FSA Mathematics is grades 3–8



Day 5

TIMSS Benchmarks

- Florida TIMSS results from the 2011 assessment
- Administered by the International Association for the Evaluation of Educational Achievement (IEA)
- Four-year assessment cycle
- Grades 4 and 8



Day 5

PISA Benchmarks

- Florida PISA results from the 2012 assessment.
- Administered by the Organization for Economic Cooperation and Development (OECD)
- Three-year cycle
- Age 15 students
- We will use PISA benchmarks for the EOC tests and Grade 10 ELA



Day 5

- Submit recommendations for Round 4



Questions?



Rooms for Each Breakout Session

	Rooms	Subject	Staff in room
1	Regency Hall 1	Algebra 1	Kari Stellpflug / Christina Estes
2	Regency Hall 2	Algebra 2	Chris Paskoff / Susan Sherwood
3	Regency Hall 9	Geometry	Sam Thomas / Marie Kristine-Tardif
4	Regency Hall 3	Grade 3 Math	Alysa Kartee / Tiffany Abu-Shaikha
5	Regency Hall 4	Grade 4 Math	Jim McCann / Lisa Schaaf
6	Regency Hall 5	Grade 5 Math	Paul Maxon / Daniel Freedberg
7	Regency Hall 6	Grade 6 Math	Erica Ajder / Eileen Heneghan
8	Regency Hall 7	Grade 7 Math	Maureen Font / Nate Thompson
9	Regency Hall 8	Grade 8 Math	Jennifer Rubel / Bernard Farley
10	Grand Cypress A	Grade 3 ELA	Allison Stingley / Stephanie Ryan
11	Grand Cypress B	Grade 4 ELA	John Neral / Jacob Wilkes
12	Grand Cypress C	Grade 5 ELA	Sean Redmond / Kevin Clayton
13	Grand Cypress G	Grade 6 ELA	Brett Craycraft / Sarah Abdelnaby
14	Grand Cypress H	Grade 7 ELA	Diana Reed / Terra Winsett
15	Grand Cypress I	Grade 8 ELA	Natalie Rebentisch / Amber Benlian
16	Poinciana AB	Grade 9 ELA	Kelly Quinney / Brian Kline
17	Poinciana CD	Grade 10 ELA	Katina Marshall / Anthony Kazanjian



www.FLDOE.org

