# Standard Setting Results for the Oklahoma Alternate Assessment Program

Dr. Michael Clark
Research Scientist
Psychometric & Research Services
Pearson

State Board of Education Meeting 28 June 2012



#### **General Overview**

- The Oklahoma Alternate Assessment Program (OAAP)
   Portfolio is part of the Oklahoma School Testing
   Program. The OAAP is the accountability program for students with severe disabilities (referred to as the 1% assessment in the No Child Left Behind Act).
- The OAAP is a portfolio-based assessment program and is a form of a criterion referenced assessment that ties directly to the domain of knowledge and skill defined by the PASS content standards.



#### **Rationale for Setting Standards in 2012**

- Standard setting is required to align the performance level expectations for students qualified to take the OAAP with the updated scoring rubrics, including realignment of evidence to the PASS content standards.
- Standards were set for all OAAP assessments (26 tests):
  - Mathematics 3-8
  - Reading 3-8
  - Writing 5 & 8
  - Science 5 & 8
  - Social Studies 5, 7, & 8
  - End-of-Instruction Algebra I, Algebra II, Geometry, Biology I, English II, English III, and U.S. History

OKLAHOMA STATE DEPARTMENT OF EDUCATION

#### **Standard Setting and Cut Scores**

- Standard setting is the process whereby experts make judgments about the content that a student should know and what he/she should be able to do in order to be classified in a specific performance level.
- Three cut scores divide the possible scores on a given assessment into the four performance levels utilized in Oklahoma:
  - 1) Unsatisfactory
  - 2) Limited Knowledge
  - 3) Proficient
  - 4) Advanced



#### **Performance Level Descriptors (Grade 3 Math)**

**Unsatisfactory:** The student at the **unsatisfactory** level will be able to sort objects by number, size, and other properties; count with 1:1 correspondence to 10; identify circles; recognize coins; and identify a table/chart.

Limited Knowledge: In addition to skills described at the lower achievement levels, the student at the limited knowledge level will be able to copy a pattern; place numbers in correct numerical order; identify circular and linear items; identify coins; and collect data for a table/chart.



#### **Performance Level Descriptors (Grade 3 Math)**

**Proficient:** In addition to skills described at the lower achievement levels, the student at the **proficient** level will be able to describe the classification system used to categorize two groups of items; demonstrate an understanding of "half" and "whole"; describe/sort circular and linear items; indicate values of coins; and organize data into a table/chart.

**Advanced:** In addition to skills described at the lower achievement levels, the student at the **advanced** level will be able to create and extend patterns; create three collections and indicate which is greatest; explain differences between circles and squares; match coins to a given amount; and interpret data in a table/chart.

OKLAHOMA STATE DEPARTMENT OF EDUCATION

Janet Barresi, State Superintendent of Public Instruction

#### **Threshold Descriptions**

- The threshold student
  - Borderline or minimally qualified in the context of a particular performance level
  - Just barely meets criteria to be included in the performance level
- Committees developed descriptions of threshold students at each performance level cut point
  - Concrete; related to the performance level descriptors
  - Committees described at least three characteristics per cut
- Threshold description discussion centered on the following qualities of threshold level students:
  - What should they do?
  - What skills should they possess?
  - What should they know?
  - What academic behaviors demonstrate that they are at a particular performance level?

#### **The Standard Setting Process**

- Standard setting committees convened in mid-June to determine cut score recommendations.
- Recommendations were based on:
  - Task specifications;
  - Scoring rubrics;
  - Evaluation of portfolios;
  - The performance level descriptors and how they related to specific portfolios.
- The cut scores were NOT based on the number of points awarded to a portfolio or the desired distribution of students across the four performance levels.



#### **Standard Setting – Methods**

- Use reasoned judgment to understand scoring structure and assign expectations of performance.
- Use body-of-work method for evaluating cut scores.
- Reasoned Judgment:
  - Students can perform differently on each task and yet earn the same total score. See table below.
  - Four portfolios as examples.

Student	Task 1	Task 2	Task 3	Task 4	Task 5	Total Score
Α	4	2	3	1	4	14
В	2	3	3	3	3	14
С	4	4	4	1	1	14
D	4	4	2	2	QKLAI	HOMA STATE

#### **Standard Setting - Methods**

- Reasoned Judgment
  - Training
    - Provide judgments for expectations regarding how students should perform on each of the tasks (content standards).
    - Use the scoring rubric to assign a value for student performance.
    - Do this for each cut score Limited Knowledge, Proficient, and Advanced.
    - Consider the threshold student at each performance level when making judgments.
    - Consider only one performance level at a time and judge across tasks, then move to next performance level.
    - All judgments were made independently.
    - Opportunities for discussion based on levels of rater agreement were presented between rounds of judgment.



#### **Standard Setting – Methods**

- Body-of-work
  - Feedback and Discussion
    - Review of portfolios after each round of judgments.
    - Portfolios provided at and around (one score point lower and one score point higher) the recommended cut scores.
    - Share thoughts on portfolios with regard to judgments on the tasks.
    - Evaluate if judgments are too lenient or too strict at each performance level.
    - Consider whether recommended cut scores align with expectations for student performance on the portfolios (i.e., threshold descriptions)



#### **Standard Setting - Methods (cont.)**

- Body-of-work Holistic Judgment
  - Final Decision
    - Review rating of committee Round Three cut scores.
    - Propose final recommendations for each cut score.
    - Holistic judgment.



#### **Vertical Articulation**

- Vertical Articulation is the evaluation of the patterns of impact data across subjects/grades based on the subject/grade-level committee recommendations.
  - Evaluate all cut scores and impact data.
- Use subject/grade-level committee members as panelists to represent their committee.
- Evaluate if expectations (i.e., similar standards of rigor) and if judgments were consistent across committees.
  - Review potential variation across committees not explained by expectations of students and the content of the tests.
    - View all subjects and all grades.



#### **Purpose - Policy Review**

- Expectation:
  - Review the cut scores and, in particular impact data, and evaluate if it makes sense from a policy perspective.
- Policy committee: used impact data to recommend potential adjustments to the cut scores.
- Panelists included
  - Business leaders
  - Key stakeholders



#### **Standard Setting Committee Members**

- The experts included on the standard setting committees were educators with significant experience in instruction at the appropriate grade level and subject with special education students.
- General education teachers were also selected for a balanced committee.
- The 8 committees consisted of over 150 experts who set standards for the 26 portfolio assessment programs.
- The committees varied in size from 15 to 20 panelists.
- Vertical Articulation
  - 21 individuals from standard setting committees
- Policy Review
  - 4 independent stakeholders



#### **Standard Setting Committee Members**

- The standard setting committees represented teachers from across the state, from districts and schools of various sizes, and were diverse in terms of race/ethnicity.
- Each committee had a wealth of classroom experience.

#### Years in Current Position

Committee	Average	Total
Mathematics 3-5	8.6	164.0
Mathematics 6-8	8.4	135.0
Reading 3-5 & Writing 5	12.8	244.0
Reading 6-8 & Writing 8	10.4	167.0
Science 5 & 8, Biology I	6.4	102.0
Social Studies 5, 7, & US History	10.3	174.5
Algebra I, II, & Geometry	10.6	169.0
English II & III	7.3	124.5

#### **Recommended Performance Benchmarks (Cut Scores)**

- The committees' recommendations for cut scores will be presented next.
- The results are provided as:
  - Raw scores that would be used for Spring 2012 as the cuts under the recommendation from these committees.
  - Impact data, which show how students that completed the assessments in Spring 2012 would be classified using the recommended cuts.
- Recommendations from the vertical articulation committee.
- Policy review committee confirmed cut scores and impact data.



#### **OAAP Mathematics 3-8: Recommended Cut Scores**

Subject	Grade	Cut Score						
		LK		Pro		Adv		
		Raw	%	Raw	%	Raw	%	
Math	3	8	40	12	60	18	90	
	4	10	42	16	67	21	88	
	5	7	35	12	60	17	85	
	6	9	38	15	63	23	96	
	7	6	30	13	65	19	95	
	8	7	35	13	65	19	95	



# OAAP Reading and Writing 3-8: Recommended Cut Scores

Subject	Grade	Cut Score						
		LK		Pro		Adv		
		Raw	%	Raw	%	Raw	%	
	. 3	6	30	12	60	18	90	
Read	4	6	30	11	55	17	85	
	5	5	31	9	56	14	88	
	Writ 5	5	25	11	55	18	90	
	6	5	31	10	63	14	88	
	7	8	33	14	58	20	83	
	8	8	33	14	58	21	88	
	Writ 8	7	44	11	69	15	94	



# OAAP Science 5 & 8 and Social Studies 5, 7, & 8: Recommended Cut Scores

Subject	Grade	Cut Score						
		LK		Pro		Adv		
	·	Raw	%	Raw	%	Raw	%	
Science	5	10	36	16	57	25	89	
	8	14	39	22	61	32	89	
Social Studies	5	13	41	20	63	29	91	
	7	8	40	12	60	18	90	
	8	9	38	15	63	22	92	

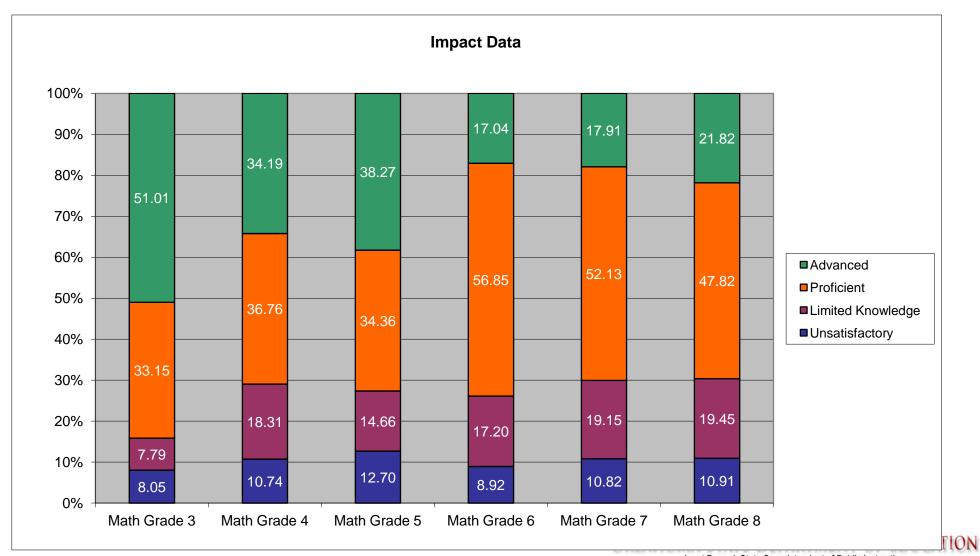


#### **OAAP End-of-Instruction: Recommended Cut Scores**

Subject	Grade	Cut Score					
		LK		Pro		Adv	
		Raw	%	Raw	%	Raw	%
Algebra I	HS	6	38	10	63	15	94
Algebra II	HS	4	33	8	67	11	92
Biology I	HS	16	40	25	63	35	88
English II	HS	14	39	22	61	31	86
English III	HS	10	36	17	61	25	89
Geometry	HS	5	31	10	63	15	94
U.S. History	HS	12	38	21	66	30	94

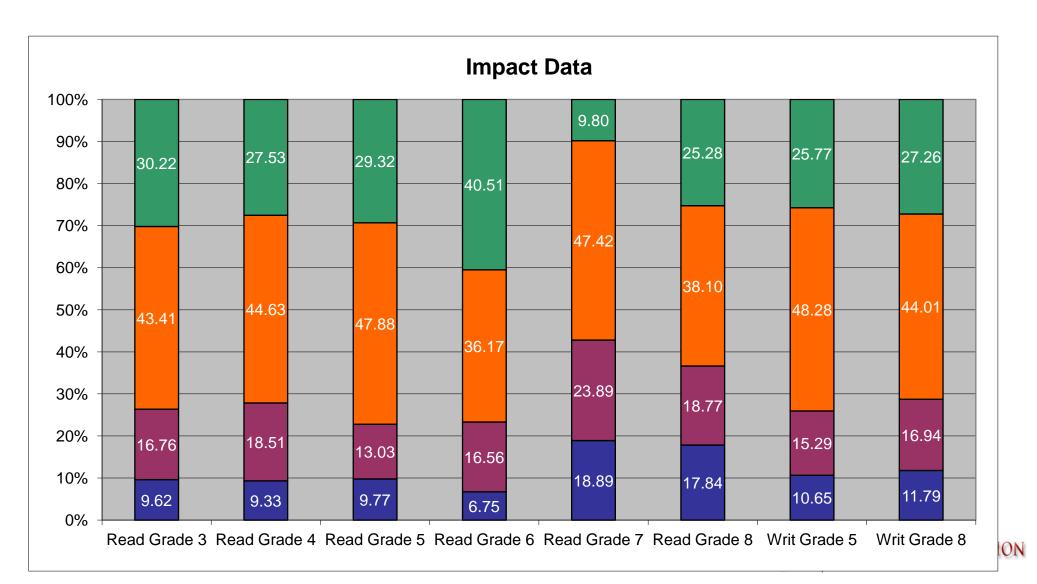


# **Impact for OAAP Mathematics 3-8 Cut Scores**

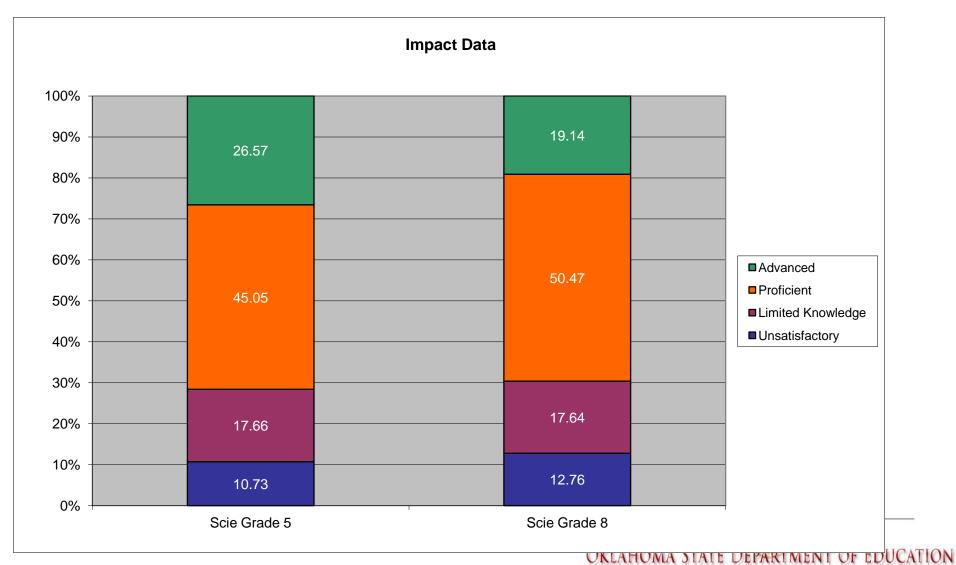


Janet Barresi, State Superintendent of Public Instruction

### **Impact for OAAP Reading & Writing 3-8 Cut Scores**

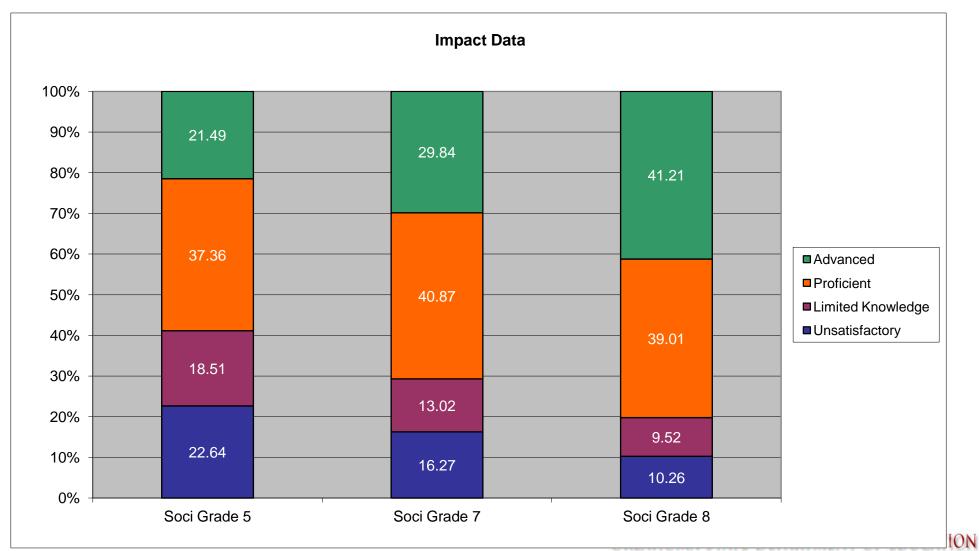


### **Impact for OAAP Science 5 & 8 Cut Scores**



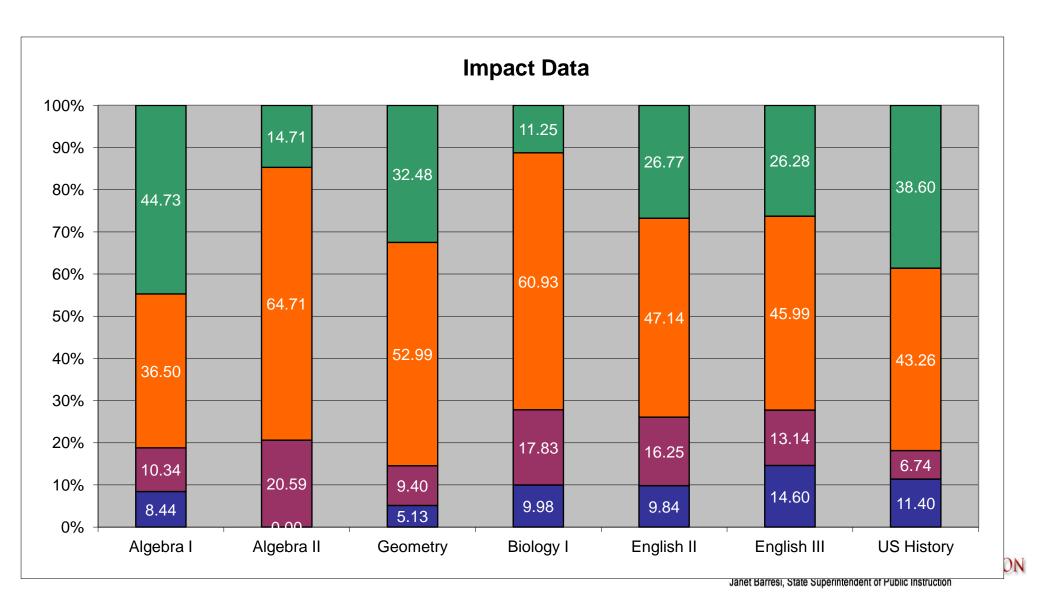
Janet Barresi, State Superintendent of Public Instruction

### **Impact for OAAP Social Studies 5, 7, & 8 Cut Scores**



Janet Barresi, State Superintendent of Public Instruction

#### **Impact for OAAP End-of-Instruction Cut Scores**



#### Recommendations

- Accept the recommended cut scores from the standard setting committees.
- Make these recommended cuts effective immediately (i.e., apply to the student scores collected in Spring 2012).



# **Thank you**

