Appendix 23

Algebra I End-of-Course Standard Setting Report

Algebra I End-of-Course (EOC) Exam
Standard-Setting

On November 3–4, 2005, the Texas Education Agency (TEA) and Pearson Educational Measurement (PEM) convened a group of 22 panelists to recommend performance standards—specifically, Met Standard and Commended Performance—for the Algebra I End-of-Course (EOC) Exam, which was administered in the current format for the first time in spring 2005. The two-day standard-setting event included sessions in which panelists (1) reviewed the test items, (2) clarified what the performance-level descriptors meant for the Algebra I course, and (3) applied an item-mapping procedure to set recommended cut scores (Lewis, Green, Mitzel, Baum, & Patz, 2003; Mitzel, Lewis, Patz, & Green, 2001). During the item-mapping procedure, panelists reviewed the content assessed by the test items, engaged in table and whole group discussions, and considered the impact on students when making their cut score recommendations. This report describes the process and outcomes of the standard-setting session.

Background

The Algebra I EOC Exam is an optional online test that assesses students' mastery of algebra after taking the Algebra I course. The assessment includes items that cover Algebra I content from the Texas Essential Knowledge and Skills (TEKS) curriculum. This version of the Algebra I EOC exam was administered online for the first time in spring 2005 to 20,858 students who had completed their Algebra I course. Score reports for this administration included only raw score information about student test performance. TEA asked PEM to convene a standard-setting advisory panel to recommend performance standards for the Algebra I EOC Exam. Standards on this exam will (1) provide students and educators with valuable information on the meaning associated with particular raw scores, (2) aid in the interpretation of test results from administration to administration, and (3) allow for an evaluation of student performance relative to the performance standards within a given administration. The new cut scores will be reported for students taking the examination during the fall testing window (November 28, 2005–February 28, 2006).

Purpose and Structure of the Standard-Setting Meeting

The purpose of the standard-setting panel was to recommend to TEA performance standards for the Algebra I EOC Exam that correspond to the performance-level descriptors for Did Not Meet Standard, Met Standard, and Commended Performance. To facilitate panelists' judgments, the session included three main activities:

1. a description of the history of the Algebra I EOC Exam, an explanation of the purpose of the standard-setting, and a review of the test questions;
2. clarification of what the performance-level descriptors mean for Algebra I; and
3. participation in a standard-setting procedure called item mapping.

Panel Composition

Twenty-two panelists participated in the standard-setting process. The gender composition of the panel was 36% male and 64% female. The ethnic composition of the panel was 14% African American, 27% Hispanic, and 59% Caucasian. When the gender and ethnic composition of the panelists was compared to the composition reported for Algebra I students who took the exam in spring 2005, males and Hispanics were slightly underrepresented on the panel (The students who took the test were 50% male, 50% female, 13% African American, 36% Hispanic, and 36% Caucasian. No ethnicity designation was provided for 13% of the students who took the test and 2% of the students reported their ethnicity as Other. Panelists from all 20 regions of the state were invited to attend. The panelists who attended the meeting represented 18 of the 20 regions. Panelists represented small, medium, and large districts from rural, suburban, and urban

communities. Panelists had a wide range of expertise, including three department chairs/heads, one math coordinator, one math specialist, two Education Service Center representatives, and one curriculum leader. Panelists reported teaching Algebra I, Algebra II, Geometry, pre-Calculus, Calculus, and pre-AP courses. Panelists' names and demographic characteristics are presented in Appendix A. of this report.

Overview of the Algebra I EOC Exam

The first activity for the panelists involved an overview of the Algebra I EOC Exam and the purpose of the standard-setting meeting. Victoria Young from TEA's Student Assessment Division summarized the history of the End-of-Course test and reviewed the test development process followed to create tests for this subject. Ms. Young explained the differences between the current EOC test and previous EOC tests, highlighting the fact that the current version was developed to align with the TAKS by assessing the five algebra objectives that are assessed on the TAKS mathematics tests. Panelists were also told that this was the first time a standard-setting panel had been convened to set standards on an EOC exam.

The Testing Interface

The Algebra I EOC Exam is administered to students as a computerized online assessment. Part of the panelists' time at the standard-setting meeting was spent on computers so that the panel could experience the test in the same way that students would. The panel viewed two online tutorials; the first was a slideshow tutorial highlighting key components of the online interface, and the second was an interactive tutorial that allowed panelists to practice using the online testing tools available for use during the test. Panelists also took the test on computer, just as students would.

Clarification of the Performance-Level Descriptions

After panelists were given the context in which they were to make recommendations about the Algebra I EOC standards, a whole-group discussion was used to translate the general performance-level descriptors into concrete student academic behaviors specific to Algebra I. For this activity, panelists described the types of skills and levels of proficiencies that students meeting each performance-level descriptor should possess. Complete results from this activity are listed in Appendix B. After concrete student academic behaviors were generated for each performance level, panelists developed a preliminary draft of the three characteristics that most differentiated Met Standard from Did Not Meet Standard and Met Standard from Commended Performance. This preliminary draft was refined and finalized during a group discussion at the end of the standard-setting meeting.

The three characteristics that panelists felt most differentiated Did Not Meet Standard and Met Standard were that students who Met Standard could

1. read and apply algebraic terminology and reasoning to problem-solving contexts;
2. use multiple strategies (including technology) in problem solving; and
3. have the diligence and organizational strategies to work problems through to completion.

The three characteristics that panelists felt most differentiated Met Standard from Commended Performance were that students who demonstrated Commended Performance could

1. evaluate the problem-solving situation to select the most effective strategies;
2. synthesize algebraic knowledge in problem solving; and
3. extend beyond the mechanics of algebra to a conceptual understanding.

The Item-Mapping Process

After panelists clarified what the performance-level descriptors meant for Algebra I, they were trained in an item-mapping standard-setting process and participated in three rounds of judgments. PEM facilitators introduced the general item-mapping process and explained the tasks required of panelists. Before making their judgments for each round, panelists were asked to read each item, identify the knowledge and skills needed for a correct response, and review the performance-level descriptors and list of student academic behaviors. As they made their judgments, panelists were asked to imagine a group of students who just barely Met Standard. In a test booklet of items ordered by difficulty from easiest to hardest (ordered item booklet), panelists were asked to identify the last item which a student who just barely Met Standard would answer correctly. Item difficulties were used to order the 50 live items in the ordered item booklet and the 10 field-test items in the practice exercise booklet. The ordered item booklet and practice exercise booklet were provided to the panelists in both hardcopy and computerized versions. For each round of ratings, panelists were instructed to view the item booklet on the computer while making their decisions. This process was designed to remind the panelists to think about the computer interface as part of the Algebra I EOC test. Item difficulties for all items (both live and field-test) were estimated using Winsteps software (Linacre & Wright, 2000). Panelists worked independently while making their judgments and were reminded to reconsider items in the "neighborhood" of their preliminary judgment before making a final decision. After identifying a cut score for Met Standard, panelists used a similar process to make judgments for the Commended Performance cut score.

After round 1, panelists reviewed anonymous information about their fellow panelists' judgments. Then they discussed in table groups their reasons for making their judgments for round 1. After this discussion PEM facilitators provided item p-values (the percent of students who answered the items correctly in the spring 2005 test administration) for all the ordered items in the test booklet. Panelists discussed these item-level data and then individually reevaluated their cut score decisions and provided their judgments for round 2.

Before round 3, panelists were again shown a summary of their fellow panelists' anonymous judgments, after which they discussed in table groups the knowledge and skills required of the items near the group median cut scores. Panelists were then given impact data (the percents of students expected to be in the three performance categories based on spring 2005 test data). The impact data included overall percents and percents by group (ethnicity, gender, and economically disadvantaged status). PEM facilitators also explained possible limitations of the test data. Panelists discussed the impact data in their table groups and then as a whole group. For round 3, panelists again made independent recommendations for each cut score.

Incorporating the Testing Interface into the Standard-Setting Process

The online testing interface provided a unique opportunity to incorporate computers into the standard-setting process. As mentioned above, the panelists completed tutorials and took the Algebra I EOC exam on the computer. In addition, the practice exercise booklet and ordered item booklet were presented on the computer. To facilitate the use of the computers during the meeting the room set-up had to allow panelists easy access to the computer and access to table space for materials and table discussions. One computer was provided for each panelist so that they could view the computerized version of the Algebra I EOC items at any time during the meeting. Additional computers were also available for use by observers of the standard setting meeting. Two rows of 12 computers (for a total of 24 computers) were set up facing each other, with an additional third row of tables between the two rows with the computers (see Appendix C for a diagram of the room set-up). Panelists could sit at the computer to work individually through the ordered item booklet and then turn their chair to be sitting at their group table for table discussions and whole group discussions. Viewing screens were placed at either end of the rows and presentation materials were projected simultaneously on both screens. This allowed panelists at either end of the room to have an

unobstructed view of the standard setting presentation. The facilitator could move up and down the rows behind the panelists to present at either end of the room. Panelists were reminded to make use of the computerized ordered item booklet throughout the meeting.

Panelists' Judgments Across Rounds

Exhibit A shows panelists' median judgments for the two cut scores across the three rounds. Panelists' median judgments increased from Round 1 to Round 2 but remained the same from round 2 to round 3 for both the Met Standard and Commended Performance cut scores. Exhibits B and C illustrate all the panelists' raw-score judgments for the Met Standard and Commended Performance ratings across the three rounds. The graphs in Exhibits B and C show the decrease in variability in panelists' judgments across rounds, with the largest decrease occurring from Round 1 to Round 2.

Exhibit A: Panelists' Median Judgments for the Met Standard and Commended Performance Cut Scores Across Rounds

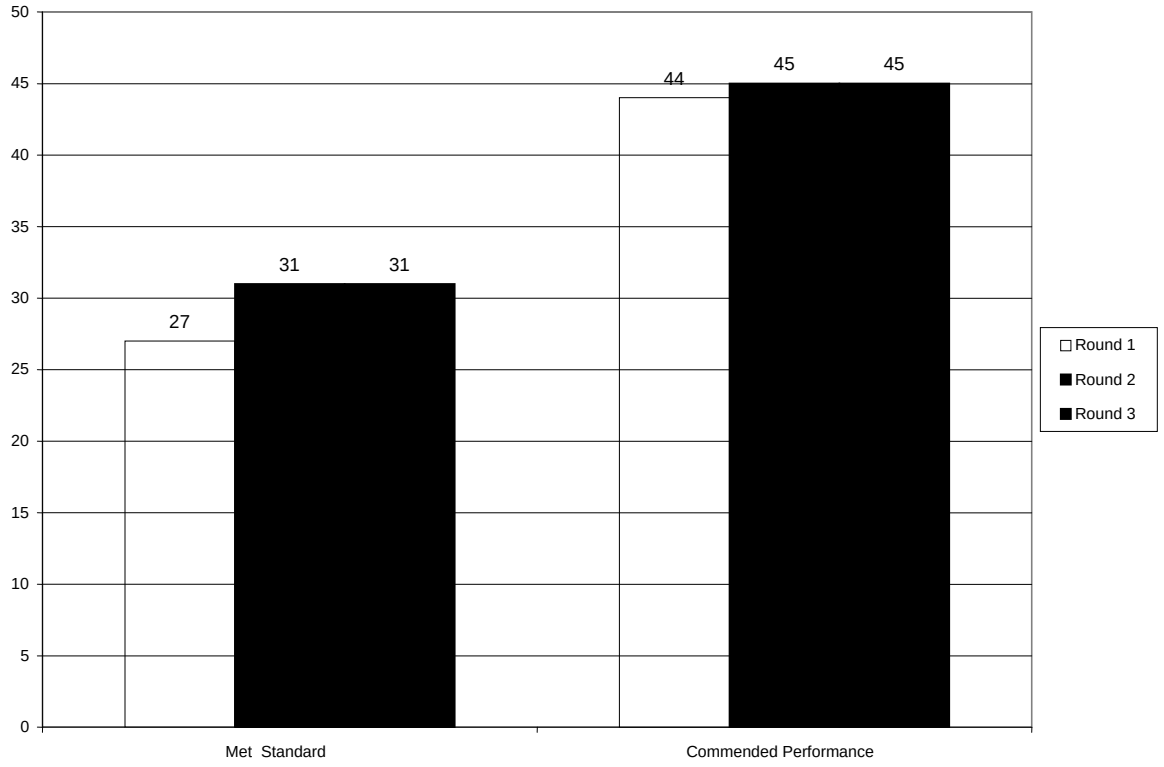


Exhibit B: Panelists' Raw-Score Judgments for the Met Standard Cut Across Rounds

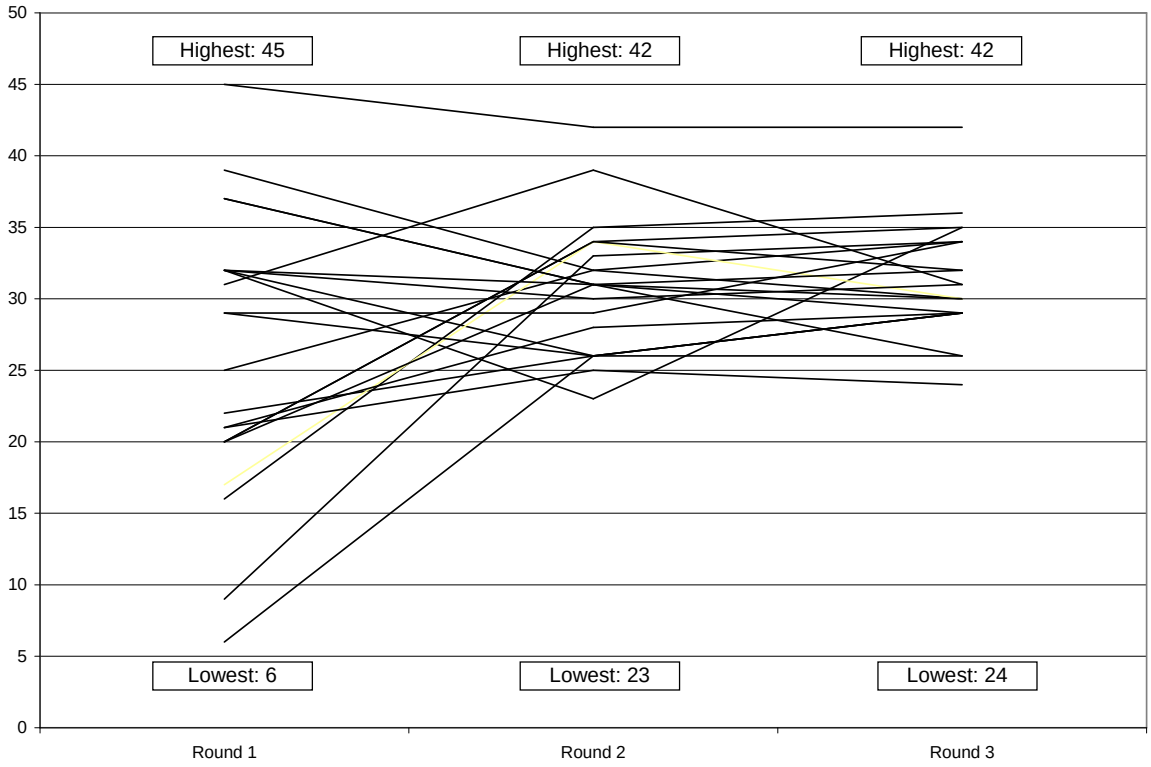


Exhibit C: Panelists' Raw-Score Judgments for the Commended Performance Cut across Rounds

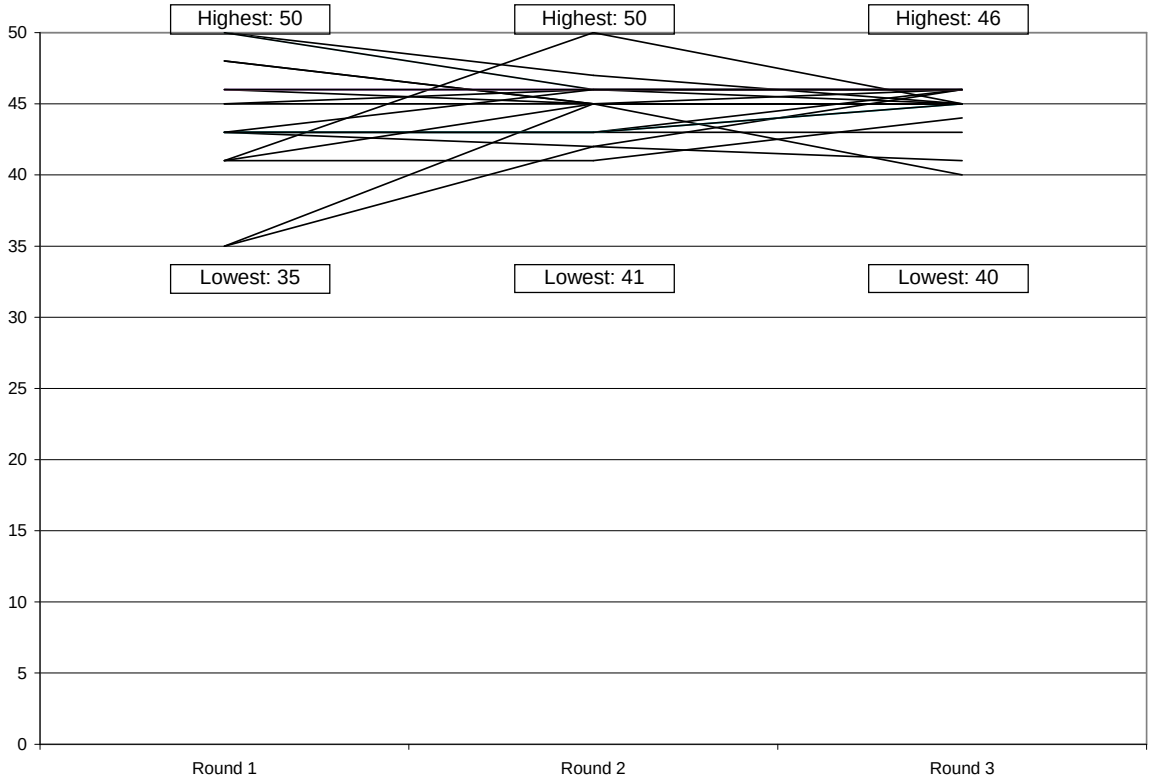


Table Groups

Discussions during the meeting were either conducted with the whole group or in table groups. Four table groups were formed with either five or six panelists in each group. Panelists were assigned to table groups before the start of the meeting to try to ensure a balanced gender, ethnic, and regional representation within each group. Table group medians for each round are displayed graphically in Exhibits D and E. Medians, means, and standard deviations for table groups are displayed in Exhibits F and G.

Exhibit D: Table Group Median Judgments for the Met Standard Cut Across Rounds

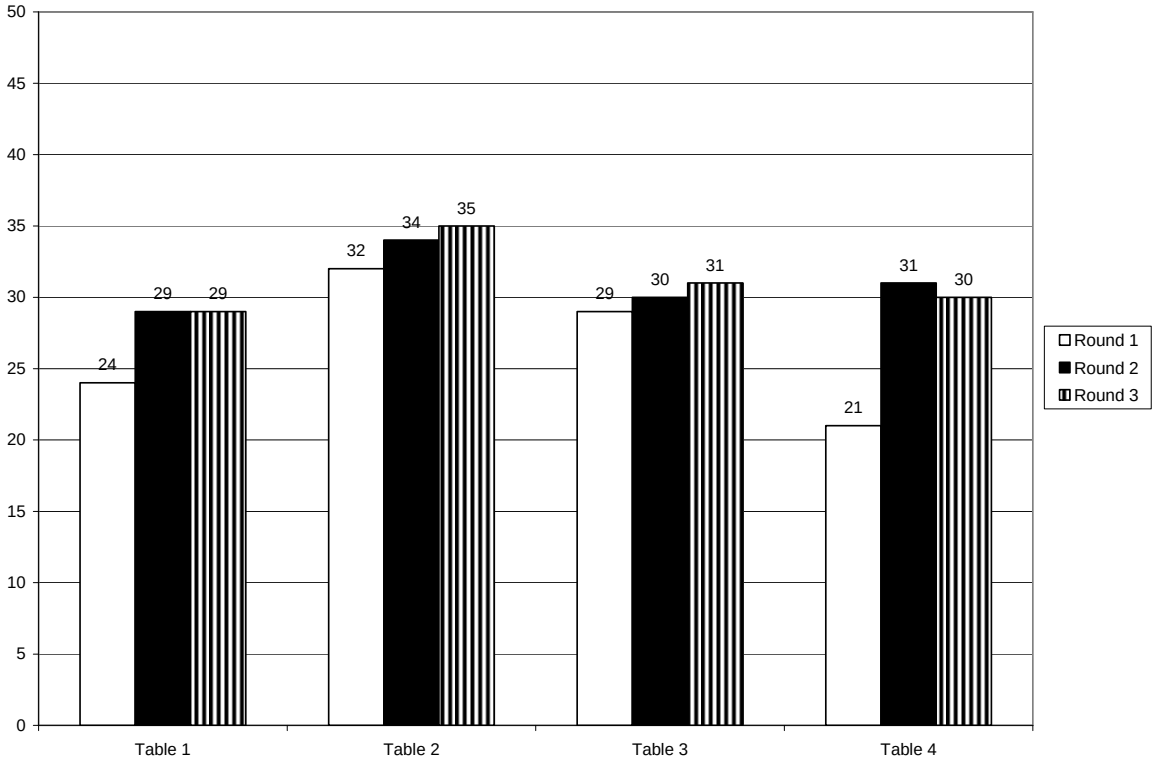


Exhibit E: Table Group Median Judgments for the Commended Performance Cut Across Rounds

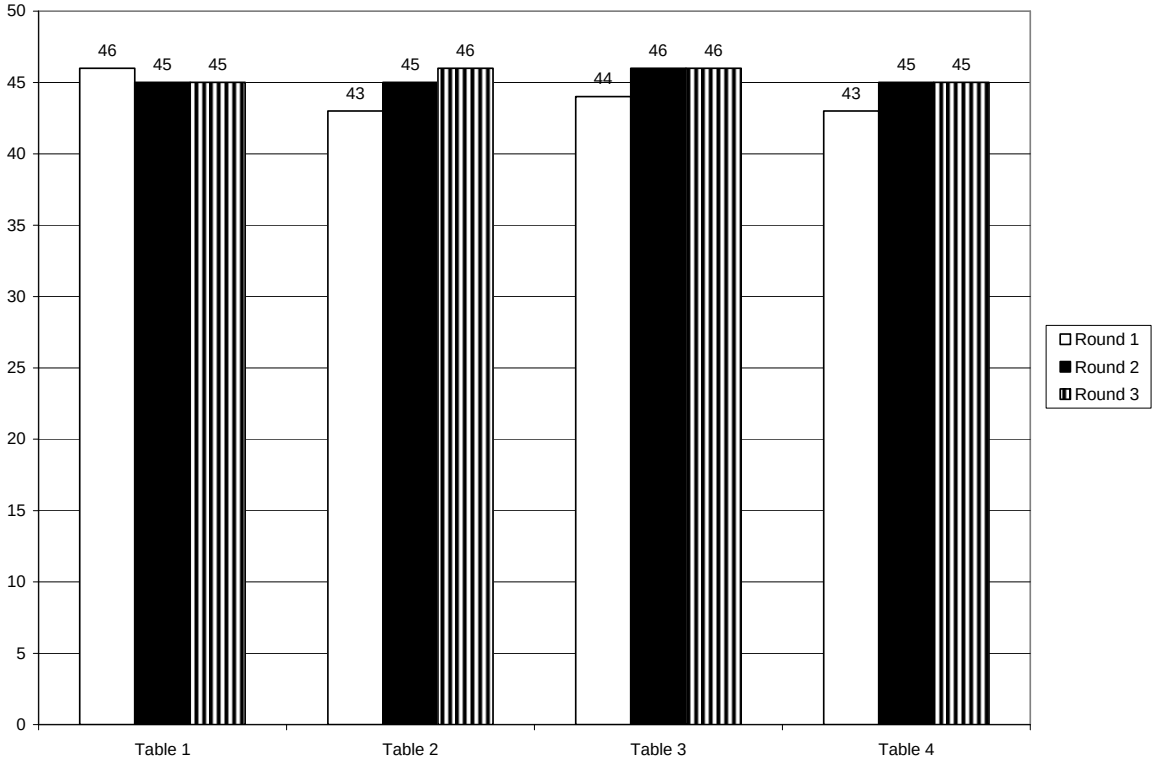


Exhibit F: Table Group Statistics for the Met Standard Cut Across Rounds

		Raw Score Cuts (Maximum Raw Score = 50)								
		Round 1			Round 2			Round 3		
	n	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD
Table 1	6	24	21.8	12.3	29	29.0	3.4	29	29.5	2.6
Table 2	5	32	30.4	12.3	34	33.2	6.8	35	36.4	3.2
Table 3	6	29	26.5	6.1	30	30.5	5.2	31	29.3	3.6
Table 4	5	21	26.0	8.0	31	31.0	2.1	30	30.4	1.5

Exhibit G: Table Group Statistics for the "Commended Performance" Cut Across Rounds

		Raw Score Cuts (Maximum Raw Score = 50)								
		Round 1			Round 2			Round 3		
	n	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD
Table 1	6	46	46.0	1.9	45	45.0	1.1	45	44.0	2.2
Table 2	5	43	44.2	6.2	45	44.6	2.1	46	45.8	0.5
Table 3	6	44	43.2	5.3	46	45.2	3.3	46	44.7	2.0
Table 4	5	43	43.0	1.4	45	44.4	1.3	45	45.2	0.5

Sample Comparison with Texas Assessment of Knowledge and Skills (TAKS)

After round 1, panelists were shown student performance data (percent of students correctly answering each item) based on the spring 2005 test administration. Additional information was provided for the panelists to allow them to compare the sample of students taking the Algebra I EOC Exam in spring 2005 with the sample of students participating in the mathematics TAKS (grades 9 and 10) during spring 2005. Since the Algebra I EOC Exam is an optional test, it was important to compare the sample in terms of gender, ethnicity, and economic status composition with the sample of students who are required to take the TAKS tests. Exhibit H displays the percentage of students taking the Algebra I EOC exam, the grade 9 TAKS mathematics test, and the grade 10 TAKS mathematics test by breakout group. The samples of students taking each type of test are similar in composition.

Exhibit H: Percentage (Number) of Students Taking Algebra I EOC, TAKS Mathematics 9, TAKS Mathematics 10 – Spring 2005

	Algebra I EOC		TAKS Mathematics 9		TAKS Mathematics 10	
Gender						
Male	50%	(10,354)	50%	(160,721)	49%	(131,830)
Female	50%	(10,490)	49%	(157,534)	50%	(134,371)
Missing	0%	(14)	0%	(380)	0%	(218)
Ethnicity						
Caucasian	36%	(7,475)	40%	(128,896)	44%	(117,385)
Hispanic	36%	(7,596)	42%	(133,081)	38%	(101,952)
African American	13%	(2,617)	14%	(45,286)	14%	(36,347)
Other	2%	(404)	3%	(10,633)	4%	(10,328)
Missing	13%	(2,766)	0%	(739)	0%	(407)
Economically Disadvantaged						
No	61%	(12,632)	54%	(172,955)	60%	(159,433)
Yes	39%	(8,226)	45%	(144,602)	40%	(106,327)
Missing	0%	(0)	0%	(1,078)	0%	(659)
Total	100%	(20,858)	100%	(31,8635)	100%	(266,419)

Impact Data

Statewide impact data by group for the final recommended cut scores are presented in Exhibit I. Impact data were calculated using live-test data from the spring 2005 testing window. Approximately 20,858 students participated in the test administration. Each student completed a single form of the online exam with the same 60 items. Of those 60 items, 50 items were live items and 10 items were embedded field-test items. Students omitting more than 5 items on the test were not included in the final calculation of the impact data, resulting in a sample of 20,033 students. Raw scores on the 50 items were used to calculate the percentage of students (overall and from different groups) within each performance level. This information was then used to provide impact data based on the panelists' recommended cut scores during the standard-setting meeting.

Exhibit I: Impact Data (Percent of Students in Each Category) for the Final Recommended Cut Scores

Group	Number of Students Used to Calculate Impact Data	Did Not Meet Standard	Met Standard	Commended Performance
Overall	20,033	49%	45%	7%
Males	9,947	50%	43%	7%
Females	10,073	47%	46%	7%
Caucasian	7,263	48%	53%	10%
African American	2,556	70%	29%	2%
Hispanic	7,131	56%	40%	4%
Other	394	29%	57%	14%
Economically Disadvantaged	7,844	61%	36%	3%

Note. The category percentages are exclusive of each other.

Calculating the Standard Errors

Two standard errors were calculated and are reported for the standard-setting meeting. The first was the standard error of the mean panelists' ratings. This standard error was calculated as the standard deviation of panelists' round 3 votes divided by the square root of the sample size (22). The standard error of the median was estimated as the standard error of the mean multiplied by the constant 1.253 (MacCann & Stanley, 2004). In addition, the estimated standard error of measurement and reliability for the test was calculated using data from the spring 2005 administration of the Algebra I EOC Exam. The estimated SEM (Test) is 3.0 and the estimated reliability (coefficient alpha) is 0.90.

Data Summary

The recommendations from the standard-setting panel and their associated standard errors are summarized in Exhibit J below.

Exhibit J: Final Recommended Cut Scores and Standard Errors

	Raw Score Cuts (Maximum Raw Score = 50)					
	Met			Commended		
	Median	Mean	SD	Median	Mean	SD
Round 1	27	26.0	9.8	44	44.1	4.1
Round 2	31	30.8	4.6	45	44.8	2.0
Round 3	**31**	31.2	4.0	**45**	44.9	1.6

	Round 3 Summary Statistics	
	Met	Commended
Number of Panelists	22	22
SE (Mean)	0.8	0.3
SE (Median)	1.1	0.4

Evaluation Survey

After the final recommendations were made by the group, all the panelists completed an evaluation of the standard-setting meeting. Nearly all the panelists reported that they thought the main session activities were successful or very successful. One panelist (5%) reported that the discussion of the performance categories and the descriptors was not successful, and two panelists (9%) reported that the overview of the item-mapping method and practice exercise was partially successful.

Panelists also reported how useful the activities and information were in helping them make recommendations. Of all responses in this section, most indicated that the activities and information were useful or very useful. Five panelists (23%) found the state-provided descriptors for each performance category somewhat useful. One panelist (5%) reported that the group discussion of the descriptors to make them more concrete was not useful and one panelist (5%) reported that the information provided by the facilitator prior to Round 1 was somewhat useful. Three panelists (14%) reported that the student

performance data and the state impact data were not useful or somewhat useful. All the panelists (100%) reported that the training, feedback, and total time allowed for ratings were adequate or more than adequate. One panelist (5%) reported that the facilities were not adequate.

Eighteen panelists (82%) reported confidence that the Round 3 judgments represented appropriate levels of student performance for the Met Standard cut score, and all the panelists (100%) reported confidence in the Commended Performance cut score. If the panelists were not confident in the Round 3 cut scores, they were asked where they would have moved the cut. Some panelists responded to this item even though they had responded that they were confident in the cut scores. None of the panelists (0%) reported that another round of ratings would have been helpful.

All twenty-two panelists (100%) reported that they were able to express their professional opinions about student performance levels, express their opinions about appropriate standards of student performance, ask questions about the uses to be made of the standards, ask questions about the process of making recommendations, and interact with fellow judges. And all the panelists (100%) reported that their opinions and judgments were treated with respect from fellow judges and moderators. Full results of the evaluation survey are presented in Appendix D.

References

Lewis, D. M., Green, D. R., Mitzel, H.C., Baum, K. & Patz, R.J. (2003). *The Bookmark procedure: Methodology and recent implementations*. Paper presented at the 1998 Annual Meeting of the National Council on Measurement in Education. San Diego, CA.

Linacre, J.M. & Wright, B.D. (2000). *A user's guide to WINSTEPS*. Chicago: MESA Press.

MacCann, R. G. & Stanley, G. (2004). Estimating the standard error of the judging in a modified-angoff standards setting procedure. *Practical Assessment, Research & Evaluation*, *9(5)*. Located at http://PAREonline.net/getvn.asp?v=9&n=5

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.

Appendix A

Panelist Information

Number	Nominee	District	ESC Region	Gender	Ethnicity	Subject(s) Taught	Grade(s)	Position
1	Arroyo, Nancy	Ysleta	19	Female	Hispanic	Math, Sci, Alg. 1, Alg. 2, Precalc, Geom.	NA	Dept. Head
2	Beaty, Tracy	Gunter	10	Male	Caucasian	Alg. 1, Geom.	9–10	Teacher
3	Carmona, Belinda	Kingsville	2	Female	Hispanic	Alg. 1, Alg. 2	HS	Teacher
4	Cosenza, Gary	ESC IV	4	Male	Caucasian	Math	HS	ESC
5	De la Pena, Daniel	Kerrville	20	Male	Hispanic	Alg. 1, Alg. 2	9–11	Teacher
6	Gerardo, Lolita	PSJA	1	Female	Hispanic	Alg. 1, Geom.	HS	Dept. Chair
7	Griffin, Joy	Pittsburg	8	Female	Caucasian	Alg. 1, Alg. 2	9–12	Teacher
8	Hebert, Larry	Kirbyville	5	Male	Caucasian	Alg. 1, Geom.	HS	Teacher
9	Hernandez, Rogelio	McAllen	1	Male	Hispanic	Alg. 1, Alg. 2, Geom.	HS	Teacher
10	Johnson, Tom	ESC XVI	16	Male	Caucasian	Math	Prev. taught 8–11 Math	ESC
11	King, Michelle	Coppell	10	Female	African American	Math	All	Math Coordinator
12	Mares, Heather	Early	15	Female	Caucasian	Alg. 1	9	Teacher
13	Mesa, Brenda	Temple	12	Female	Caucasian	Alg. 1, Pre-AP/IB Alg. 2, Geom., Pre-AP Geom.	9–12	Master Math Teacher, Curriculum Leader for Alg. 1
14	Payne, Vic	Fabens	19	Male	Caucasian	Alg. 1, Precalc	NA	Dept. Chair Head
15	Ramirez, Jane	Wichita Falls	9	Female	Hispanic	Pre-AP Alg. 1, Geom.	9–12	Teacher
16	Roberson, Jr., Uris	Mineola	7	Male	African American	Alg. 1, Alg. 2, Geom.	MS–HS	Teacher
17	Roberts, Holly	Midland	18	Female	Caucasian	Math	K–12	Math Director
18	Schwertner, Trina	Plainview	17	Female	Caucasian	Alg. 1, Math Plus	6, 9	Teacher
19	Sisco Snow, Karen	Austin	13	Female	Caucasian	Alg, 1, Geom., Alg. 2, Precalc, Calc, Math	HS	Math Specialist
20	Thomas, Dale	Houston	4	Female	African American	Alg. 1	9–11	Teacher
21	Wood, Lynda	Arlington	11	Female	Caucasian	Alg. 1, Geom.	9–10	Teacher
22	Wood, Stephanie	Victoria	3	Female	Caucasian	Alg. 2, Precalc, Geom.	9–12	Teacher

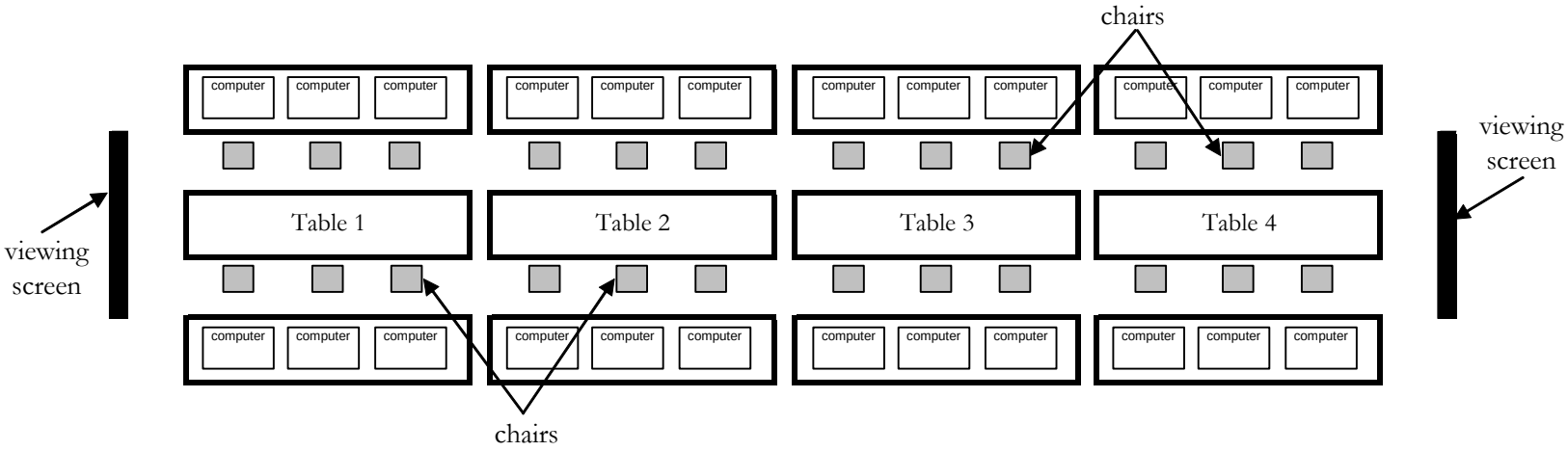
Appendix B

Panelists' Concrete Examples of Student Academic Behaviors in Algebra I for Performance-Level Descriptors

Did Not Meet Standard	Met Standard	Commended Performance
Can use a graphing calculator but without understanding the concept	Can use graphing calculator as a tool	Can program graphing calculators to do what they want them to do
Do not recognize wrong answers as being unreasonable	Can eliminate unreasonable answer choices	Can learn from mistakes
Can complete 1- or 2 step problems but often accept the first answer they get (giving up easily)	Can approach multi-step problems requiring reading	Are able to make conjectures
Are not motivated to achieve	Are motivated to achieve/persevere and persist	Can excel through motivation to overcome skill deficiencies
Have no concept of what a variable represents	Understand the concept of a variable	Can identify examples and counterexamples
Have difficulty transferring math skills to real-life settings	Can transfer knowledge	Can transfer knowledge within math across other subjects
Have weak higher-order thinking skills	Are able to use abstract reasoning (concrete↔abstract)	Have good deductive reasoning skills; can reason from the inverse
Lack prior knowledge of mathematics terminology (many gaps)	Have basic math skills and vocabulary	Has a conceptual (not just mechanical) understanding of multiple representations of data/relationships
Can do "naked math" but not in context (such as word problems)	Can make connections between patterns and algebraic representations	Can make links between related concepts (Pythagorean Theorem vs. Distance Formula)
Have little or no retention	Can retain ideas	Can restate ideas and articulate algebraic concepts
Cannot verbalize how to solve a problem	Can reason algebraically	Can create new knowledge, synthesize information, and go beyond the problem
Can do basic operations but do not know the order of operations	Can navigate between different representations (verbal, symbolic, graphic)	Have basic skills that have become second nature and are used reflexively
See little relevance for mathematics in their own lives and lack confidence in mathematics	Are confident in mathematics and will ask questions (seeking clarification and making connections)	Can focus on critical, important issues; not limited to the little details
Usually read below grade level/ do not complete assignments/ lack responsibility for own behavior/ have short attention span.	Possess test-taking skills/ can read with comprehension /take the time to read test questions carefully	Are highly motivated/ enjoy a challenge/ enjoy math/ have leadership skills/ are competitive/ have goals
Cannot apply strategies (know them but cannot use them)	Know multiple ways to approach mathematics problems	Can practice multiple methods in problem solving and checking answers
Can do patterns but cannot generalize	Are good with patterns, have an understanding and knowledge of graphs/tables/charts	
Have poor study skills/organizational skills	Have organizational skills	
	Can select and use strategies to solve problems	

Appendix C

Room Set-Up to Allow Computer Access While Setting Standards



Appendix D

Evaluation Survey Results

Standard-Setting Advisory Panel
Algebra I End-of-Course Exam
November 2005

The purpose of this evaluation form is to collect information about your experiences in recommending student performance standards for Algebra I. Your opinions provide an important part of our evaluation of the training procedures and the standard-setting method and results.

Please do not write your name or judge number on this Evaluation Form as we want your views to be anonymous. Thank you for your opinions and time.

1. Check the column below that best reflects your opinion about the level of success of these various components of the session in which you have just participated. The activities should have helped you both understand the process and feel confident in the judgments you made.

	Not Successful	*Partially Successful*	*Successful*	*Very Successful*
a. Introduction to the process of setting performance standards	0 (0%)	0 (0%)	8 (36%)	14 (64%)
b. Discussion of the performance categories and the descriptors	1 (5%)	0 (0%)	5 (23%)	16 (73%)
c. Experiencing taking the actual Algebra I EOC test	0 (0%)	0 (0%)	8 (36%)	14 (64%)
d. Overview of the item-mapping method and practice exercise	0 (0%)	2 (9%)	6 (27%)	14 (64%)
e. Feedback on judges' ratings following Round 1	0 (0%)	0 (0%)	8 (36%)	14 (64%)
f. Group discussions following Round 1	0 (0%)	0 (0%)	7 (32%)	15 (68%)
g. Discussions following Round 2	0 (0%)	0 (0%)	5 (23%)	17 (77%)

2. How useful do you feel the following activities or information were in assisting you to make your recommendations?

	Not Useful	*Somewhat Useful*	*Useful*	*Very Useful*	*Omit*
a. State-provided descriptors for each performance category	0 (0%)	5 (23%)	8 (36%)	9 (41%)	0 (0%)
b. Group discussion of the descriptors to make them more concrete	1 (5%)	0 (0%)	6 (27%)	15 (68%)	0 (0%)
c. Information provided by the facilitator prior to Round 1	0 (0%)	1 (5%)	8 (36%)	12 (55%)	0 (0%)

	Not Useful	*Somewhat Useful*	*Useful*	*Very Useful*	*Omit*
d. Feedback provided concerning your panel's judgments after Round 1 (both peer ratings and discussion)	0 (0%)	0 (0%)	5 (23%)	17 (77%)	0 (0%)
e. Student item-performance data provided for Rounds 2 and 3	1 (5%)	2 (9%)	4 (18%)	14 (64%)	1 (5%)
f. Feedback provided concerning your panel's judgments after Round 2 (peer ratings and discussion)	0 (0%)	0 (0%)	9 (41%)	13 (59%)	0 (0%)
g. Projected "state impact data" provided prior to Round 3	2 (9%)	1 (5%)	6 (27%)	13 (59%)	0 (0%)

3. How adequate were the following elements of the session?

	Not Adequate	*Somewhat Adequate*	*Adequate*	*More Than Adequate*
a. **Training** provided prior to Round 1	0 (0%)	0 (0%)	8 (36%)	14 (64%)
b. Amount of **time** spent training before Round 1	0 (0%)	0 (0%)	8 (36%)	14 (64%)
c. **Feedback** provided between rounds	0 (0%)	0 (0%)	4 (18%)	18 (82%)
d. **Facilities** used for the sessions	1 (5%)	0 (0%)	8 (36%)	13 (59%)
e. **Total amount of time** for the ratings	0 (0%)	0 (0%)	10 (45%)	12 (56%)

4. In applying the standard-setting method, your panel was asked to recommend two cut scores for student performance on Algebra I EOC. How confident do you feel that the **descriptions of the categories** are reasonable for each student performance level?

	Not Confident	*Somewhat Confident*	*Confident*	*Very Confident*
a. Met Standard	1 (5%)	2 (9%)	8 (36%)	11 (50%)
b. Commended Performance	0 (0%)	0 (0%)	9 (41%)	13 (59%)

5. With regard to the final Round 3 ratings:

a. How confident do you feel that your Round 3 ratings represent appropriate levels of student performance for the Met Standard cut score?

Not Very Confident	*Confident*	*Very Confident*	*Omit*
3 (14%)	5 (23%)	13 (59%)	1 (5%)

b. If you are not very confident with the final recommended placement of the cut score for Met Standard, would you move the placement before or after the current page?

Before	*After*	*NA*	*Omit*
1 (5%)	4 (18%)	14 (64%)	3 (1%)

c. If you checked before or after, how many pages would you move in that direction?

1 Page After	*4 Pages After*	*2 Pages Before*	*NA*	*Omit*
1 (5%)	3 (14%)	1 (5%)	14 (64%)	3 (14%)

d. How confident do you feel that your Round 3 ratings represent appropriate levels of student performance for the Commended Performance cut score?

Not Very Confident	*Confident*	*Very Confident*	*Omit*
0 (0%)	4 (18%)	17 (77%)	1 (5%)

e. If you are not confident with the final recommended placement of the cut score for Commended Performance, would you move the placement before or after the current page?

Before	*After*	*NA*	*Omit*
0 (0%)	4 (18%)	13 (59%)	5 (23%)

f. If you checked before or after, how many pages would you move in that direction?

1 Page After	*2 Pages After*	*NA*	*Omit*
3 (14%)	1 (5%)	13 (59%)	5 (23%)

6. Do you think you would have benefited from an additional round of ratings?

Another round would have been helpful.	*Three rounds were sufficient*
0 (0%)	22 (100%)

7. Did you have adequate opportunities during the session to:

	Not Adequate	*Adequate*
a. Express your professional opinions about student performance levels?	0 (0%)	22 (100%)
b. Express your professional opinions about appropriate standard of student performance?	0 (0%)	22 (100%)

Did you have adequate opportunities during the session to:

	Not Adequate	*Adequate*
c. Ask questions about the uses to be made of the standards?	0 (0%)	22 (100%)
d. Ask questions about the process of making recommendations?	0 (0%)	22 (100%)
e. Interact with your fellow judges on your views?	0 (0%)	22 (100%)

8. Do you believe your opinions and judgments were treated with respect by your fellow judges and by the moderators?

No	*I'm not sure*	*Yes*
0 (0%)	0 (0%)	22 (100%)

9. Please use the space below to make any additional comments you wish to make about the process or your experience. Thank you for taking the time to evaluate the session.

- I have learned very much about this process and hope to be invited to do similar tasks in the future. Thanks for this opportunity to look at the assessment.
- It was very professionally done. Thank you.
- I commend you for doing an excellent job.
- I believe that the performance-level descriptors were helpful but not the most significant reason for setting my cut scores. I relied heavily on my past experiences in the classroom.
- This gave me a better understanding of how parameters are made.
- This was a very worthwhile experience. The facilitators were very knowledgeable and provided very detailed resources in a timely manner. The facilities were more than adequate and a great deal of planning made it possible. The level of thinking and participation by all members of the advisory panel was great. I feel the task at hand was completed, resulting in a high-quality final product. Thank you for the opportunity.
- Enjoyed the experience.
- Use a square for large group discussion. A long rectangle makes is difficult for everyone to participate. Re-arrange the room for better viewing of screen.
- Thanks for this opportunity. Have you considered taping these sessions as sharing with non-participants and SBOE as part of the presentation and recommendations? Also please inform the panel when the recommendations go to the SBOE so that if we want to attend we can.
- It was a great experience. I enjoyed working with such a professional and dedicated group of teachers. I learned a lot about what is happening at other districts.
- Extremely useful data-generating system. I never understood the levels of understanding that are needed to generate commended and met commended students.
- Thought three rounds were sufficient, but we should have changed groups.
- I have enjoyed this experience very much. I have a new respect for the process and the agencies involved, and would welcome the opportunity to work on another panel or with either agency in some capacity. This is a very professionally run project, and everyone was very helpful and friendly. Thank you for this great experience and I would welcome another.
- In your online program. The rule tool should have the point of rotation to be to "0" on the ruler. Not the end of the ruler. A better rule would allow the student to click on the center of the ruler to drag the whole thing around. And on one of the ends to rotate it about the rotation point. The chart screen should open as minimized, however, it should have the option to max it so you do not have to scroll up or down. As a "pic" it should be set to fit the window when maximized. Please use the scroll wheel on the mouse. The "back," "next," "reset," "go to," and "review" buttons need to be at the top with the tools buttons, not at the bottom where you must scroll to see them in some questions.