Listening. Learning. Leading.

Multistate Standard-Setting Technical Report

PRAXIS™ MATHEMATICS: CONTENT KNOWLEDGE (5161)

Licensure and Credentialing Research

ETS

Princeton, New Jersey

March 2013

Copyright © 2013 by Educational Testing Service. All rights reserved. ETS, the ETS logo, and LISTENING. LEARNING. LEADING. are registered trademarks of Educational Testing Service (ETS) in the United States of America and other countries throughout the world. Praxis is a trademark of ETS.

EXECUTIVE SUMMARY

To support the decision-making process of education agencies establishing a passing score (cut score) for the Praxis™ Mathematics: Content Knowledge (5161) test, research staff from Educational Testing Service (ETS) designed and conducted a multistate standard-setting study.

PARTICIPATING STATES

Panelists from 24 states and Washington, DC were recommended by their respective education agency. The education agencies recommended panelists with (a) experience either as mathematics teachers or college faculty who prepare mathematics teachers and (b) familiarity with the knowledge and skills required of beginning mathematics teachers.

RECOMMENDED PASSING SCORE

ETS provides a recommended passing score from the multistate standard-setting study to help education agencies determine an appropriate operational passing score. For the Praxis Mathematics: Content Knowledge test, the recommended passing score[1] is 32 out of a possible 50 raw-score points. The scaled score associated with a raw score of 32 is 160 on a 100–200 scale.

[1] Results from the two panels participating in the study were averaged to produce the recommended passing score.

To support the decision-making process for education agencies establishing a passing score (cut score) for the Praxis™ Mathematics: Content Knowledge (5161) test, research staff from ETS designed and conducted a multistate standard-setting study in February 2013 in Princeton, New Jersey. Education agencies[2] recommended panelists with (a) experience, either as mathematics teachers or college faculty who prepare mathematics teachers and (b) familiarity with the knowledge and skills required of beginning mathematics teachers. Twenty-four states and Washington, DC(see Table 1) were represented by 35 panelists. (See Appendix A for the names and affiliations of the panelists.)

Table 1

Participating Jurisdictions and Number of Panelists

Alaska (1 panelist)	North Dakota (1 panelist)
Arkansas (2 panelists)	Pennsylvania (1 panelist)
Delaware (1 panelist)	Rhode Island (1 panelist)
Idaho (2 panelists)	South Carolina (1 panelist)
Kentucky (1 panelist)	South Dakota (1 panelist)
Louisiana (1 panelist)	Tennessee (1 panelist)
Maine (1 panelist)	Utah (2 panelists)
Maryland (2 panelists)	Vermont (1 panelist)
Mississippi (2 panelists)	Washington, DC (1 panelist)
Nevada (1 panelist)	Wisconsin (2 panelists)
New Hampshire (2 panelists)	West Virginia (1 panelist)
New Jersey (2 panelists)	Wyoming (2 panelists)
North Carolina (2 panelists)	

The following technical report contains three sections. The first section describes the content and format of the test. The second section describes the standard-setting processes and methods. The third section presents the results of the standard-setting study.

ETS provides a recommended passing score from the multistate standard-setting study to education agencies. In each jurisdiction, the department of education, the board of education, or a designated educator licensure board is responsible for establishing the operational passing score in

[2] States and jurisdictions that currently use Praxis were invited to participate in the multistate standard-setting study.

accordance with applicable regulations. This study provides a recommended passing score,[3] which represents the combined judgments of two panels of experienced educators. Each jurisdiction may want to consider the recommended passing score but also other sources of information when setting the final Praxis Mathematics: Content Knowledge passing score (see Geisinger & McCormick, 2010). A jurisdiction may accept the recommended passing score, adjust the score upward to reflect more stringent expectations, or adjust the score downward to reflect more lenient expectations. There is no *correct* decision; the appropriateness of any adjustment may only be evaluated in terms of its meeting the jurisdiction's needs.

Two sources of information to consider when setting the passing score are the standard error of measurement (SEM) and the standard error of judgment (SEJ). The former addresses the reliability of the Praxis Mathematics: Content Knowledge test score and the latter, the reliability of panelists' passing-score recommendation. The SEM allows a jurisdiction to recognize that any test score on any standardized test—including a Praxis Mathematics: Content Knowledge test score—is not perfectly reliable. A test score only *approximates* what a candidate truly knows or truly can do on the test. The SEM, therefore, addresses the question: How close of an approximation is the test score to the *true* score? The SEJ allow a jurisdiction to gauge the likelihood that the recommended passing score from a particular panel would be similar to the passing scores recommended by other panels of experts similar in composition and experience. The smaller the SEJ, the more likely that another panel would recommend a passing score consistent with the recommended passing score. The larger the SEJ, the less likely the recommended passing score would be reproduced by another panel.

In addition to measurement error metrics (e.g., SEM, SEJ), each jurisdiction should consider the likelihood of classification error. That is, when adjusting a passing score, policymakers should consider whether it is more important to minimize a false-positive decision or to minimize a false-negative decision. A false-positive decision occurs when a candidate's test score suggests he should receive a license/certificate, but his actual level of knowledge/skills indicates otherwise (i.e., the candidate does not possess the required knowledge/skills). A false-negative decision occurs when a candidate's test score suggests that she should not receive a license/certificate, but she actually does possess the required

[3] In addition to the recommended passing score <u>averaged</u> across the two panels, the recommened passing scores for <u>each</u> panel are presented.

knowledge/skills. The jurisdiction needs to consider which decision error may be more important to minimize.

Overview of the Praxis Mathematics: Content Knowledge Test

The Praxis Mathematics: Content Knowledge *Test at a Glance* document (ETS, in press) describes the purpose and structure of the test. In brief, the test measures whether entry-level mathematics teachers have the knowledge/skills believed necessary for competent professional practice.

The two-hour assessment contains 60 selected-response and numeric-entry items[4] covering two content areas: *Number and Quantity, Algebra, Functions, and Calculus* (approximately 41 items) and *Geometry, Probability and Statistics, and Discrete Mathematics* (approximately 19 items).[5] The reporting scale for the Praxis Mathematics: Content Knowledge test ranges from 100 to 200 scaled-score points.

Processes and Methods

The design of the standard-setting study included two, independent expert panels. Before the study, panelists received an email explaining the purpose of the standard-setting study and requesting that they review the content specifications for the test. This review helped familiarize the panelists with the general structure and content of the test.

For each panel, the standard-setting study began with a welcome and introduction by the meeting facilitator. The facilitator described the test, provided an overview of standard setting, and presented the agenda for the study. Appendix B shows the agenda for the panel meeting.

[4] Ten of the 60 selected-response and numeric-entry items are pretest items and do not contribute to a candidate's score.
[5] The number of items for each content area may vary slightly from form to form of the test.

Reviewing the Test

The standard-setting panelists first took the test and then discussed it. This discussion helped bring the panelists to a shared understanding of what the test does and does not cover, which serves to reduce potential judgment errors later in the standard-setting process.

The test discussion covered the major content areas being addressed by the test. Panelists were asked to remark on any content areas that would be particularly challenging for entry-level teachers or areas that address content particularly important for entry-level teachers.

Defining the Target Candidate

Following the review of the test, panelists described the target candidate. The *target candidate description* plays a central role in standard setting (Perie, 2008); the goal of the standard-setting process is to identify the test score that aligns with this description.

Panel 1 created a description of the target candidate — the knowledge/skills that differentiate a *just* from a *not quite* qualified candidate. To create this description, the panel first split into smaller groups to consider the target candidate. The full panel then reconvened and, through whole-group discussion, created the description of the target candidate to use for the remainder of the study.

The written description of the target candidate summarized the panel discussion in a bulleted format. The description was not intended to describe all the knowledge and skills of the target candidate but only highlight those that differentiate a *just* qualified candidate from a *not quite* qualified candidate. The written description was distributed to panelists to use during later phases of the study (see Appendix C for the target candidate description).

For Panel 2, the panelists began with the description of the target candidate developed by Panel 1. Given that the multistate standard-setting study was designed to provide two recommendations for the same performance standard, it was important that panels use consistent target candidate description to frame their judgments. The panelists reviewed the target candidate description, and any ambiguities were discussed and clarified.

PANELISTS' JUDGMENTS

The standard-setting process for the Praxis Mathematics: Content Knowledge test was a probability-based Modified Angoff method (Brandon, 2004; Hambleton & Pitoniak, 2006). In this study, each panelist judged each item on the likelihood (probability or chance) that the target candidate would answer the item correctly. Panelists made their judgments using the following rating scale: 0, .05, .10, .20, .30, .40, .50, .60, .70, .80, .90, .95, 1. The lower the value, the less likely it is that the target candidate would answer the item correctly because the item is difficult for the target candidate. The higher the value, the more likely it is that the target candidate would answer the item correctly.

Panelists were asked to approach the judgment process in two stages. First, they reviewed both the description of the target candidate and the item and decided if, overall, the item would be difficult for the target candidate, easy for the target candidate or moderately difficult/easy. The facilitator encouraged the panelists to consider the following rules of thumb to guide their decision:

- Difficult items for the target candidate are in the 0 to .30 range.
- Moderately difficult/easy items for the target candidate are in the .40 to .60 range.
- Easy items for the target candidate are in the .70 to 1 range.

Next, panelists decided how to refine their judgment within the range. For example, if a panelist thought that an item would be easy for the target candidate, the initial decision located the item in the .70 to 1 range. The second decision for the panelist was to decide if the likelihood of answering it correctly is .70, .80, .90, .95 or 1.

After the training, panelists made practice judgments and discussed those judgments and their rationale. All panelists completed a post-training survey to confirm that they had received adequate training and felt prepared to continue; the standard-setting process continued only if all panelists confirmed their readiness.

Following this first round of judgments (*Round 1*), item-level feedback was provided to the panel. The panelists' judgments were displayed for each item and summarized across panelists. Items were highlighted to show when panelists converged in their judgments (at least two-thirds of the panelists located an item in the same difficulty range) or diverged in their judgments.

The panelists discussed their item-level judgments. These discussions helped panelists maintain a shared understanding of the knowledge/skills of the target candidate and helped to clarify aspects of

items that might not have been clear to all panelists during the Round 1 judgments. The purpose of the discussion was not to encourage panelists to conform to another's judgment, but to understand the different relevant perspectives among the panelists.

In Round 2, panelists discussed their Round 1 judgments and were encouraged by the facilitator (a) to share the rationales for their judgments and (b) to consider their judgments in light of the rationales provided by the other panelists. Panelists recorded their Round 2 judgments only for items when they wished to change a Round 1 judgment. Panelists final judgments for the study, therefore, consist of their Round 1 judgments and any adjusted judgments made during Round 2.

Other than the description of the target candidate, results from Panel 1 were not shared with Panel 2. The item-level judgments and resulting discussions for Panel 2 were independent of judgments and discussions that occurred with Panel 1.

RESULTS

EXPERT PANELS

Table 2 presents a summary of the panelists' demographic information. The panel included 35 educators representing 24 states and Washington, DC. (See Appendix A for a listing of panelists.) Twenty-four panelists were teachers, nine were college faculty, one was an administrator or department head, and one held another position. All nine of the faculty members' job responsibilities included the training of mathematics teachers.

The number of experts by panel and their demographic information are presented in Appendix D (see Table D1).

Table 2

Panel Member Demographics (Across Panels)

	N	%
Current position		
Teacher	24	69%
Administrator/Department head	1	3%
College faculty	9	26%
Other	1	3%
Race		
White	29	83%
Black or African American	2	6%
Hispanic or Latino	3	9%
Native Hawaiian or Other Pacific Islander	1	3%
Gender		
Female	20	57%
Male	15	43%
Are you currently certified to teach this subject in your state?		
Yes	29	83%
No	6	17%
Are you currently teaching this subject in your state?		
Yes	31	89%
No	4	11%
Are you currently supervising or mentoring other teachers of this subject?		
Yes	18	51%
No	17	49%
At what K–12 grade level are you currently teaching this subject?		
Middle school (6–8 or 7–9)	1	3%
High school (9–12 or 10–12)	20	57%
Middle and High school	3	9%
Not currently teaching at the K–12 level	11	31%

Table 2 (continued)
Panel Member Demographics (Across Panels)

	N	%
Including this year, how many years of experience do you have teaching this subject?		
3 years or less	4	11%
4–7 years	13	37%
8–11 years	5	14%
12–15 years	6	17%
16 years or more	7	20%
Which best describes the location of your K–12 school?		
Urban	7	20%
Suburban	9	26%
Rural	10	29%
Not currently working at the K–12 level	9	26%
If you are college faculty, are you currently involved in the training/preparation of teacher candidates in this subject?		
Yes	9	26%
No	0	0%
Not college faculty	26	74%

STANDARD-SETTING JUDGMENTS

Table 3 summarizes the standard-setting judgments (Round 2) of panelists. The table also includes estimates of the measurement error associated with the judgments: the standard deviation of the mean and the standard error of judgment (SEJ). The SEJ is one way of estimating the reliability or consistency of a panel's standard-setting judgments.[6] It indicates how likely it would be for several other panels of educators similar in makeup, experience, and standard-setting training to the current panel to recommend the same passing score on the same form of the subtest. The confidence intervals created by adding/subtracting two SEJs to each panel's recommended passing score overlap, indicating that they may be comparable.

Panelist-level results, for Rounds 1 and 2, are presented in Appendix D (see Table D2).

[6] An SEJ assumes that panelists are randomly selected and that standard-setting judgments are independent. It is seldom the case that panelists are randomly sampled, and only the first round of judgments may be considered independent. The SEJ, therefore, likely underestimates the uncertainty of passing scores (Tannenbaum & Katz, 2013).

Table 3
Summary of Round 2 Standard-setting Judgments

	Panel 1	Panel 2
Average	33.10	29.81
Lowest	20.40	27.05
Highest	45.60	34.70
SD	6.20	2.19
SEJ	1.46	0.53

Round 1 judgments are made without discussion among the panelists. The most variability in judgments, therefore, is typically present in the first round. Round 2 judgments, however, are informed by panel discussion; thus, it is common to see a decrease both in the standard deviation and SEJ. This decrease — indicating convergence among the panelists' judgments — was observed for Panel 2 (see Table D2 in Appendix D). The standard deviation increased slightly between rounds for Panel 1. The Round 2 average score is the panel's recommended passing score.

The panels' passing score recommendations for the Praxis Mathematics: Content Knowledge test are 33.10 for Panel 1 and 29.81 for Panel 2 (out of a possible 50 raw-score points). The values were rounded to the next highest whole number, to determine the functional recommended passing score — 34 for Panel 1 and 30 for Panel 2. The scaled scores associated with 34 and 30 raw points are 165 and 155, respectively.

In addition to the recommended passing score for each panel, the average passing score across the two panels is provided to help education agencies determine an appropriate passing score. The panels' average passing score recommendation for the Praxis Mathematics: Content Knowledge test is 31.46 (out of a possible 50 raw-score points). The value was rounded to 32 (next highest raw score) to determine the functional recommended passing score. The scaled score associated with 32 raw points is 160.

Table 4 presents the estimated conditional standard error of measurement (CSEM) around the recommended passing score. A standard error represents the uncertainty associated with a test score. The scaled scores associated with one and two CSEMs above and below the recommended passing score are provided. The conditional standard error of measurement provided is an estimate.

Table 4

Passing Scores Within 1 and 2 CSEMs of the Recommended Passing Score[7]

Recommended passing score (CSEM)		Scale score equivalent
32 (3.43)		160
-2 CSEMs	26	145
-1 CSEM	29	152
+ 1 CSEM	36	170
+ 2 CSEMs	39	178

Note. **CSEM = conditional standard error of measurement.**

FINAL EVALUATIONS

The panelists completed an evaluation at the conclusion of their standard-setting study. The evaluation asked the panelists to provide feedback about the quality of the standard-setting implementation and the factors that influenced their decisions. The responses to the evaluation provided evidence of the validity of the standard-setting process, and, as a result, evidence of the reasonableness of the recommended passing score.

Panelists were also shown the panel's recommended passing score and asked (a) how comfortable they are with the recommended passing score and (b) if they think the score was too high, too low, or about right. A summary of the final evaluation results is presented in Appendix D.

All panelists *strongly agreed* or *agreed* that they understood the purpose of the study. Thirty-four of the 35 panelists *strongly agreed* or *agreed* that the facilitator's instructions and explanations were clear. All panelists *strongly agreed* or *agreed* that they were prepared to make their standard-setting judgments. All but one of the panelists *strongly agreed* or *agreed* that the standard-setting process was easy to follow.

All panelists reported that the description of the target candidate was at least *somewhat influential* in guiding their standard-setting judgments; 30 of the 35 panelists indicated the description was *very influential*. All of the panelists reported that between-round discussions were at least *somewhat influential* in guiding their judgments. Slightly less than half of the panelists (14 of the 35 panelists) indicated that their own professional experience was *very influential* in guiding their judgments.

[7] The unrounded CSEM value is added to or subtracted from the rounded passing-score recommendation. The resulting values are rounded up to the next-highest whole number and the rounded values are converted to scaled scores.

All but two of the panelists indicated they were at least *somewhat comfortable* with the passing score they recommended; 23 of the 35 panelists were *very comfortable*. Thirty of the 35 panelists indicated the recommended passing score was *about right* with two of the remaining panelists indicated that the passing score was *too low* and three indicating that the passing score was *too high*.

SUMMARY

To support the decision-making process for education agencies establishing a passing score (cut score) for the Praxis Mathematics: Content Knowledge test, research staff from ETS designed and conducted a multistate standard-setting study.

ETS provides a recommended passing score from the multistate standard-setting study to help education agencies determine an appropriate operational passing score. For the Praxis Mathematics: Content Knowledge test, the recommended passing score[8] is 32 out of a possible 50 raw-score points. The scaled score associated with a raw score of 32 is 160 on a 100–200 scale.

[8] Results from the two panels participating in the study were averaged to produce the recommended passing score.

REFERENCES

Brandon, P. R. (2004). Conclusions about frequently studied modified Angoff standard-setting topics. *Applied Measurement in Education, 17*, 59–88.

ETS. (in press). *The Praxis Series: Mathematics: Content Knowledge (5161)*. Princeton, NJ: Author.

Geisinger, K. F., & McCormick, C. M. (2010), Adopting cut scores: post-standard-setting panel considerations for decision makers. *Educational Measurement: Issues and Practice, 29*, 38–44.

Hambleton, R. K., & Pitoniak, M. J. (2006). Setting performance standards. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 433–470). Westport, CT: American Council on Education/Praeger.

Perie, M. (2008). A guide to understanding and developing performance-level descriptors. *Educational Measurement: Issues and Practice, 27*, 15–29.

Tannenbaum, R. J., & Katz, I. R. (2013). Standard setting. In K. F. Geisinger (Ed.), *APA handbook of testing and assessment in psychology: Vol. 3. Testing and assessment in school psychology and education* (pp. 455–477). Washington, DC: American Psychological Association.

APPENDIX A

PANELISTS' NAMES & AFFILIATIONS

Participating Panelists With Affiliation

Panelist	**Affiliation**
Bertine Bahige	Campbell County School District #1 (WY)
Mary Bell	Peoples Academy (VT)
Lindsey Brewer	Huron High School (SD)
Sara Brown	Brown Deer High School (WI)
Emma Chandler	Salt Lake Center for Science Education (UT)
Tina Childress	Langdon Area High School (ND)
Kira Christensen	Sioux Falls Public Schools (SD)
Michelle Cirillo	University of Delaware (DE)
Jennifer Cribbs	Western Kentucky University (KY)
Nancy Jarger Daly	Millburn High School (NJ)
Brian DeMayo	Bohemia Manor High School – Cecil County Public Schools (MD)
Kimberly L. Dickerson	Southern University at New Orleans (LA)
Katy Witt Edgar	New Hope High School (SD)
Ariane Eicke	Laramie High School (WY)
Michael Fish	University of Maine at Machias (ME)
Brian Fleischer	Nevada Virtual Academy (NV)
Joseph Gonzales	Terry Sanford High School (NC)
Shiloh A. Harder	Conway High School (AR)
Jayne Heath-Wilmarth	Council Jr./Sr. High School (ID)
Christopher Hoyt	Dunbar SHS (DC)
Barry Kolar	University High School (WV)

Participating Panelists With Affiliation

Panelist	**Affiliation**
Trent Kull	Winthrop University (SC)
Meghan Leeming	Beacon Charter High School for the Arts (RI)
Deborah L. MacCullough	Cairn University (PA)
Patrice Marquette	Hollis Brookline Middle School (NH)
Ricardo Martinez	Henderson County Early College (NC)
Denise Raynes Pitcher	Western Governors University (UT)
Lincoln Robertson	White Mountains Community College (NH)
Kimberly Scarbrough	Riverside High School (AR)
Daniel M. Seaton	University of Maryland Eastern Shore (MD)
Alice Steimle	University of Mississippi (MS)
Michael Tamblyn	Whitewater High School (WI)
Christian Tomona	Newark Public Schools (NJ)
David Williams	Tennessee Department of Education (TN)
Holly Wood	Vallivue High School (ID)

APPENDIX B

STUDY AGENDA

AGENDA

Praxis Mathematics: Content Knowledge (5161) Standard-Setting Study

Day 1

Welcome and Introduction

Overview of Standard Setting & the Praxis Mathematics Test

"Take" the Praxis Mathematics Test
(Take breaks as needed)

Discuss the Praxis Mathematics Test

Define the Knowledge/Skills of a Target Candidate

Lunch

Define the Knowledge/Skills of a Target Candidate (continued)

Break

Standard Setting Training

Round 1 Standard Setting Judgments

Collect Materials; End of Day 1

AGENDA

Praxis Mathematics: Content Knowledge (5161) Standard-Setting Study

Day 2

Overview of Day 2 & Review of Training

Round 1 Feedback & Round 2 Judgments

Break

Round 1 Feedback & Round 2 Judgments (continued)

Lunch

Feedback on Round 2 Recommended Passing Score

Complete Final Evaluation

Collect Materials; End of Study

APPENDIX C

TARGET CANDIDATE DESCRIPTION

Description of the Target Candidate[9]

A target candidate …

Numbers & Quantity

1. Knows the structure and the basic operations of the various number systems with irrational and imaginary numbers
2. Knows how to work with complex numbers when solving polynomial equations and rewriting polynomial expressions
3. Knows how to determine the reasonableness of solutions within the context of a given problem
4. Understands ratios and proportions including inversely proportional relationships between two quantities

Algebra

5. Understands how to justify the reasoning process used to solve equations, including analysis of potential extraneous solutions
6. Knows how to find real and imaginary roots of a cubic

Functions

7. Understands how new functions are obtained from existing functions (e.g., domain, range, compositions, transformations, and inverses)
8. Understands how periodic phenomena are modeled using trigonometric functions
9. Understands how function behavior is analyzed using non-algebraic representations (e.g., graphs, mapping, and tables)
10. Understands how to solve trigonometric, logarithmic, and exponential equations

Calculus

11. Knows the meaning of a limit of a function (e.g., find limit from a graph)
12. Understands the derivative as a slope of a tangent line and as a rate of change
13. Knows how to approximate or evaluate derivatives and integrals numerically given a table of values or a graph
14. Understands the relationship between differentiations and integration, including the role of the fundamental theorems of calculus

[9] Description of the target candidate focuses on the knowledge/skills that differentiate a *just* from a *not quite* qualified candidate.

Description of the Target Candidate (continued)[10]

A target candidate …

Geometry

15. Understands how trigonometry is applied to non-right triangles
16. Understands arc length and area measurements of sectors of circles
17. Understands means for proving geometric properties (e.g., lines, angles, polygons, and their operations) using geometric and algebraic methods
18. Knows means for visualizing and reasoning algebraically among common 2D and 3D figures (e.g., prisms, pyramids, and cones)

Probability & Stats

19. Understands how to interpret a linear regression model (e.g., rate of change, intercepts, and correlation coefficient) in the context of the data
20. Understands and compute the concepts of interdependence and conditional probability (such as simple events, probabilities of compound events, conditional probabilities) and how to apply those concepts to data

Discrete Mathematics

21. Use logic to evaluate the truth or equivalence of statements
22. Knows how to represent arithmetic, recursive, geometric sequences and phenomena
23. Can identify and use concepts of basic set theory
24. Uses counting techniques such as permutations and combinations.

[10] Description of the target candidate focuses on the knowledge/skills that differentiate a *just* from a *not quite* qualified candidate.

APPENDIX D

RESULTS

Table D1

Panel Member Demographics (Across Panels)

	Panel 1		Panel 2	
	N	**%**	***N***	**%**
Current position				
Teacher	14	78%	10	59%
Administrator or Department Head	0	0%	1	6%
College faculty	4	22%	5	29%
Other	0	0%	1	6%
Race				
White	15	83%	14	82%
Black or African American	0	0%	2	12%
Hispanic or Latino	2	11%	1	6%
Native Hawaiian or Other Pacific Islander	1	6%	0	0%
Gender				
Female	9	50%	11	65%
Male	9	50%	6	35%
Are you currently certified to teach this subject in your state?				
Yes	16	89%	13	76%
No	2	11%	4	24%
Are you currently teaching this subject in your state?				
Yes	16	89%	15	88%
No	2	11%	2	12%
Are you currently supervising or mentoring other teachers of this subject?				
Yes	8	44%	10	59%
No	10	56%	7	41%
At what K–12 grade level are you currently teaching this subject?				
Middle school (6–8 or 7–9)	1	6%	0	0%
High school (9–12 or 10–12)	12	67%	7	41%
Middle and High school	0	0%	3	18%
Not currently teaching at the K–12 level	5	28%	7	41%

Table D1 (continued)

Panel Member Demographics (By Panel)

	Panel 1		Panel 2	
	N	**%**	***N***	**%**
Including this year, how many years of experience do you have teaching this subject?				
3 years or less	4	22%	0	0%
4–7 years	8	44%	5	29%
8–11 years	0	0%	5	29%
12–15 years	2	11%	4	24%
16 years or more	4	22%	3	18%
Which best describes the location of your K–12 school?				
Urban	4	22%	3	18%
Suburban	6	33%	3	18%
Rural	4	22%	6	35%
Not currently working at the K–12 level	4	22%	5	29%
If you are college faculty, are you currently involved in the training/preparation of teacher candidates in this subject?				
Yes	4	22%	5	29%
No	0	0%	0	0%
Not college faculty	14	78%	12	71%

Table D2

Passing Score Summary by Round of Judgments

	Panel 1		Panel 2	
Panelist	**Round 1**	**Round 2**	**Round 1**	**Round 2**
1	37.15	36.75	30.05	29.55
2	32.40	36.10	32.90	34.70
3	37.60	40.20	27.55	28.55
4	28.05	32.80	29.60	29.00
5	30.40	31.40	30.70	29.70
6	27.25	29.45	30.55	30.85
7	31.15	34.95	26.95	29.75
8	19.40	22.20	29.45	31.85
9	31.00	33.00	29.90	29.10
10	25.25	30.15	28.15	27.45
11	31.65	33.95	32.10	32.45
12	38.90	45.60	25.10	28.55
13	26.85	28.65	26.65	27.05
14	43.10	42.60	26.20	28.10
15	26.55	32.65	28.10	28.70
16	29.15	33.75	27.95	27.75
17	22.15	20.40	34.95	33.65
18	30.00	31.20		
Average	30.44	33.10	29.23	29.81
Lowest	19.40	20.40	25.10	27.05
Highest	43.10	45.60	34.95	34.70
SD	5.93	6.20	2.56	2.19
SEJ	1.40	1.46	0.62	0.53

Table D3

Final Evaluation: Panel 1

	Strongly agree		Agree		Disagree		Strongly disagree	
	N	**%**	***N***	**%**	***N***	**%**	***N***	**%**
• I understood the purpose of this study.	16	89%	2	11%	0	0%	0	0%
• The instructions and explanations provided by the facilitators were clear.	8	44%	9	50%	1	6%	0	0%
• The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	11	61%	7	39%	0	0%	0	0%
• The explanation of how the recommended passing score is computed was clear.	10	56%	8	44%	0	0%	0	0%
• The opportunity for feedback and discussion between rounds was helpful.	14	78%	3	17%	1	6%	0	0%
• The process of making the standard-setting judgments was easy to follow.	5	28%	12	67%	1	6%	0	0%

Table D3 (continued)

Final Evaluation: Panel 1

How influential was each of the following factors in guiding your standard-setting judgments?	Very influential		Somewhat influential		Not influential	
	N	%	*N*	%	*N*	%
• The description of the target candidate	14	78%	4	22%	0	0%
• The between-round discussions	14	78%	4	22%	0	0%
• The knowledge/skills required to answer each test item	12	67%	6	33%	0	0%
• The passing scores of other panel members	1	6%	13	72%	4	22%
• My own professional experience	9	50%	8	44%	1	6%

	Very comfortable		Somewhat comfortable		Somewhat uncomfortable		Very uncomfortable	
	N	%	*N*	%	*N*	%	*N*	%
• Overall, how comfortable are you with the panel's recommended passing score?	9	50%	7	39%	1	6%	1	6%

	Too low		About right		Too high	
	N	%	*N*	%	*N*	%
• Overall, the recommended passing score is:	1	6%	14	78%	3	17%

Table D4

Final Evaluation: Panel 2

	Strongly agree		Agree		Disagree		Strongly disagree	
	N	**%**	***N***	**%**	***N***	**%**	***N***	**%**
• I understood the purpose of this study.	15	88%	2	12%	0	0%	0	0%
• The instructions and explanations provided by the facilitators were clear.	14	82%	3	18%	0	0%	0	0%
• The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	9	53%	8	47%	0	0%	0	0%
• The explanation of how the recommended passing score is computed was clear.	14	82%	3	18%	0	0%	0	0%
• The opportunity for feedback and discussion between rounds was helpful.	15	88%	2	12%	0	0%	0	0%
• The process of making the standard-setting judgments was easy to follow.	9	53%	8	47%	0	0%	0	0%

Table D4 (continued)

Final Evaluation: Panel 2

How influential was each of the following factors in guiding your standard-setting judgments?	Very influential		Somewhat influential		Not influential	
	N	%	*N*	%	*N*	%
• The description of the target candidate	16	94%	1	6%	0	0%
• The between-round discussions	8	47%	9	53%	0	0%
• The knowledge/skills required to answer each test item	12	71%	5	29%	0	0%
• The passing scores of other panel members	1	6%	12	71%	4	24%
• My own professional experience	5	29%	12	71%	0	0%

	Very comfortable		Somewhat comfortable		Somewhat uncomfortable		Very uncomfortable	
	N	%	*N*	%	*N*	%	*N*	%
• Overall, how comfortable are you with the panel's recommended passing score?	14	82%	3	18%	0	0%	0	0%

	Too low		About right		Too high	
	N	%	*N*	%	*N*	%
• Overall, the recommended passing score is:	1	6%	16	94%	0	0%