Multistate Standard-Setting Technical Report

PRAXIS® MIDDLE SCHOOL MATHEMATICS (5164)

Licensure and Credentialing Research

ETS

Princeton, New Jersey

February 2021

EXECUTIVE SUMMARY

To support the decision-making process of education agencies establishing a passing score (cut score) for the *Praxis*® Middle School Mathematics (5164) test, research staff from Educational Testing Service (ETS) designed and conducted a distance-based multistate standard-setting study.

PARTICIPATING STATES

Panelists from 12 states and Washington, D.C., were recommended by their respective education agencies. The education agencies recommended panelists with (a) experience as either middle school mathematics teachers or college faculty who prepare middle school mathematics teachers and (b) familiarity with the knowledge and skills required of beginning middle school 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* Middle School Mathematics test, the recommended passing score[1] is 38 out of a possible 60 raw-score points. The scale score associated with a raw score of 38 is 157 on a 100–200 scale.

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

INTRODUCTION

To support the decision-making process for education agencies establishing a passing score (cut score) for the *Praxis*® Middle School Mathematics (5164) test, research staff from ETS designed and conducted a distance-based multistate standard-setting study in January 2021. Education agencies[2] recommended panelists with (a) experience as either middle school mathematics teachers or college faculty who prepare middle school mathematics teachers and (b) familiarity with the knowledge and skills required of beginning middle school mathematics teachers. Twelve states and Washington, D.C., were represented by 29 panelists, as listed in Table 1. (See Appendix A for the names and affiliations of the panelists.)

Table 1
Participating States, Washington, D.C., and Number of Panelists

Alabama (2 panelists)	Rhode Island (3 panelists)
Delaware (3 panelists)	South Carolina (2 panelists)
Idaho (3 panelists)	South Dakota (2 panelists)
Indiana (2 panelists)	Tennessee (2 panelists)
Kansas (2 panelists)	Washington, D.C. (2 panelists)
Kentucky (1 panelist)	West Virginia (2 panelists)
Mississippi (3 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 state and D.C., the department of education, the board of education, or a designated educator licensure board is responsible for establishing the operational passing score in accordance with applicable regulations. This study provides a recommended passing score,[3] which represents the combined judgments of two panels of experienced educators. Each state and D.C., may want to consider the recommended passing score but also other sources of information when setting the final *Praxis* Middle School Mathematics passing score (see Geisinger & McCormick, 2010). A state and

[2] States and jurisdictions that currently use *Praxis* tests were invited to participate in the multistate standard-setting study.
[3] In addition to the recommended passing score averaged across the two panels, the passing scores for each panel are presented.

D.C., 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 state and D.C.'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* Middle School Mathematics test score and the latter, the reliability of panelists' passing-score recommendation. The SEM allows a state and D.C., to recognize that any test score on any standardized test—including a *Praxis* Middle School Mathematics 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 allows a state and D.C. 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 state and D.C. should consider the likelihood of classification errors. 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 that 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 knowledge/skills. The state and D.C. need to consider which decision error is more important to minimize.

OVERVIEW OF THE *PRAXIS*® MIDDLE SCHOOL MATHEMATICS TEST

The Praxis® Middle School Mathematics *Study Companion* document (ETS, in press) describes the purpose and structure of the test. In brief, the test measures knowledge and competencies that are important for safe and effective beginning practice as a middle school mathematics teacher.

The three-hour assessment contains 66 selected-response items[4] covering five content areas: *Numbers and Operations* (approximately 16 items), *Algebra* (approximately 15 items), *Functions* (approximately 11 items), *Geometry and Measurement* (approximately 13 items), and *Statistics and Probability* (approximately 11 items).[5] The reporting scale for the *Praxis* Middle School Mathematics test ranges from 100 to 200 scale-score points.

PROCESSES AND METHODS

The design of the standard-setting study included two, independent expert panels of educators with experience with the test content and with new teachers or teacher candidates. Before the study, panelists received an email explaining the purpose of the standard-setting study and requesting that they review materials for the study, such as the test specifications and an overview presentation. This review helped familiarize the panelists with the general structure and content of the test. Additionally, panelists were asked to attend a brief, technology check meeting, to ensure that everyone could access the technology needed for the study.

For each panel, the first day of the standard-setting study began with a welcome by the meeting facilitator. After introductions of the panelists and ETS staff, the facilitator engaged the panel in a question and answer period about the overview presentation. Appendix B shows the agenda for the panel meeting.

[4] Six of the 66 selected-response 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

Test familiarization was the first activity for the panel. The purpose of test familiarization is for the panelists to review the test and become familiar with the manner in which a candidate would take the test. After the facilitator described the purpose of the review and how to access the test[6], the standard-setting panelists took the test and had a discussion of the content measured. This discussion helped bring the panelists to a shared understanding of what the test measures.

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. Overall, this discussion serves to reduce potential judgment errors later in the standard-setting process.

Defining the Just Qualified Candidate

Following the review of the test, panelists described the just qualified candidate. The *just qualified 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.

Both panels worked together to create the final description of the just qualified candidate — the knowledge/skills that differentiate a *just* from a *not quite* qualified candidate. Each panel first worked separately by working in smaller and then a large group. Then both panels convened and, through whole-group discussion, combined the two descriptions in to the final version of the just qualified candidate to use for the remainder of the study.

The written description of the just qualified candidate summarized the panel discussion in a bulleted format. The description was not intended to describe all the knowledge and skills of the just qualified 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 just qualified candidate description).

[6] The computer-administered test items were available through the ETS IBIS Content Review Tool.

PANELISTS' JUDGMENTS

The standard-setting process for the *Praxis* Middle School Mathematics 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 just qualified 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 just qualified candidate would answer the item correctly because the item is difficult for the just qualified candidate. The higher the value, the more likely it is that the just qualified 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 just qualified candidate and the item and determined what was the probability that the just qualified candidate would answer the question correctly. The facilitator encouraged the panelists to consider the following rules of thumb to guide their decision:

- Items in the 0 to .30 range were those the just qualified candidate would have a low chance of answering correctly.
- Items in the .40 to .60 range were those the just qualified candidate would have a moderate chance of answering correctly.
- Items in the .70 to 1 range were those that the just qualified candidate would have a high chance of answering correctly.

Next, panelists decided how to refine their judgment within the range. For example, if a panelist thought that there was a high chance that the just qualified candidate would answer the question correctly, the initial decision would be in the .70 to 1 range. The second decision for the panelist was to judge 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 rationales. All panelists completed a post-training evaluation 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 just qualified 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 just qualified 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 29 educators representing 12 states and D.C. (See Appendix A for a listing of panelists.) Fourteen panelists were teachers, two were mathematics instructional coaches, ten were college faculty, two were administrators or department heads, and two held other positions. All of the faculty members' job responsibilities included the training, supervising, or mentoring of middle school mathematics teachers.

The number of experts by panel and their demographic information are presented in Appendix D (Table D1). One panelist was college faculty and a department head. Another panelist was college faculty and a coach (though not specifically described as a mathematics instructional coach). As such, those numbers will not sum to 29 and the percentages will exceed 100.

Table 2

Panel Member Demographics (Across Panels)

	N	%
Current position		
Teacher	14	48
Mathematics Instructional Coach	2	7
Administrator/Department head	2	7
College faculty	10	34
Other	2	7
Race		
White	24	83
Black or African American	3	10
American Indian or Alaskan Native	1	3
Middle Eastern	1	3
Gender		
Female	23	79
Male	6	21
Are you currently certified to teach middle school mathematics in your state?		
Yes	24	83
No	5	17
Are you currently teaching middle school mathematics in your state?		
Yes	14	48
No	15	52
Are you currently supervising or mentoring other teachers of middle school mathematics?		
Yes	15	52
No	14	48
At what K–12 grade level are you currently teaching middle school mathematics?		
Elementary and Middle school	1	3
Middle school (6–8 or 7–9)	11	38
Middle and High school	1	3
High school (9–12 or 10–12)	1	3
All Grades	1	3
Other	1	3
Not currently teaching at the K–12 level	13	45

(continues on next page)

Table 2 (continued)

Panel Member Demographics (Across Panels)

	N	%
Including this year, how many years of experience do you have teaching middle school mathematics?		
3 years or less	4	14
4–7 years	4	14
8–11 years	10	34
12–15 years	4	14
16 years or more	7	24
Which best describes the location of your K–12 school?		
Urban	6	21
Suburban	5	17
Rural	5	17
Not currently working at the K–12 level	13	45
If you are college faculty, are you currently involved in the training/preparation of teacher candidates in this subject?		
Yes	8	28
No	2	7
Not college faculty	19	66

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.[7] 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 test. 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 (Table D2).

[7] 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	37.96	36.28
Lowest	33.65	28.10
Highest	43.85	40.45
SD	3.57	3.06
SEJ	0.92	0.82

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. The Round 2 average score is the panel's recommended passing score.

The panels' passing score recommendations for the *Praxis* Middle School Mathematics test are 37.96 for Panel 1 and 36.28 for Panel 2 (out of a possible 60 raw-score points). The values were rounded to the next highest whole number, to determine the functional recommended passing score — 38 for Panel 1 and 37 for Panel 2. The scale scores associated with 38 and 37 raw points are 157 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* Middle School Mathematics test is 37.12 (out of a possible 60 raw-score points). The value was rounded to 38 (next highest raw score) to determine the functional recommended passing score. The scale score associated with 38 raw points is 157.

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 scale scores associated with one and two CSEM 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 CSEM of the Recommended Passing Score[8]

Recommended passing score (CSEM)		Scale score equivalent
38 (3.76)		157
-2 CSEM	31	141
-1 CSEM	35	150
+ 1 CSEM	42	166
+ 2 CSEM	46	175

Note. **CSEM = conditional standard error(s) 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 and 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 panelists *strongly agreed* or *agreed* that the standard-setting process was easy to follow.

All panelists reported that the description of the just qualified candidate was at least *somewhat influential* in guiding their standard-setting judgments; 27 of the 29 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. More than half of the panelists (25 of the 29 panelists) indicated that their own professional experience was *very influential* in guiding their judgments.

All of the panelists indicated they were at least *somewhat comfortable* with the passing score they recommended; 24 of the 29 panelists were *very comfortable*. Twenty-seven of the 29 panelists indicated the recommended passing score was *about right* and two indicated that the passing score was *too low*.

[8] 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 scale scores.

SUMMARY

To support the decision-making process for education agencies establishing a passing score (cut score) for the *Praxis* Middle School Mathematics test, research staff from ETS designed and conducted a distance-based 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* Middle School Mathematics test, the recommended passing score[9] is 38 out of a possible 60 raw-score points. The scale score associated with a raw score of 38 is 157 on a 100–200 scale.

[9] 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®: The Praxis Study Companion: Middle School Mathematics (5164).* 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**
Ahmad Alhammouri	Jacksonville State University (AL)
Sandra Ammons	South Carolina Department of Education (SC)
Cathy Boutin	RI Math Teachers Association; Association of Teachers of Mathematics in New England (RI)
Nichole Bowman	Georgia Morse Middle School (SD)
Jeneva Clark	University of Tennessee (TN)
Beth Costner	Winthrop University (SC)
Darin DeNeal	Pendleton Heights Middle School (IN)
Ashley Digmann	Dakota Wesleyan University (SD)
Sherita Flake	Urban Teachers (DC)
Tekeeta Funchess	Hinds County Schools (MS)
Tracy Graham	North Providence School Department (RI)
Tiffany Hackendorn	Indian River School District (DE)
Shauna Hedgepeth	Purvis Middle School (MS)
Jennifer Heitman	Kamiah Middle School (ID)
Shannon Henderson	Putnam County Schools (WV)
Paul Johanson	Brigham young University - Idaho (ID)
Karen Lindsey	Germantown Middle School (MS)
Robin Magruder	Campbellsville University (KY)
Michelle Northshield	Red Clay Consolidated School District (DE)
William Reaves Jr.	Capital School District (DE)

(table continues)

Participating Panelists With Affiliation (continued)

Panelist	**Affiliation**
Terry Reed	West Virginia State University (WV)
Keri Richburg	Troy University (AL)
Diane Rodriguez	Bloomfield Jr./Sr. High School (IN)
Debra Scarpelli	Rhode Island Department of Education/Pawtucket School Department (RI)
Janet Stramel	Fort Hays State University (KS)
Jackie Vogel	Austin Peay State University (TN)
Kristopher Wallaert	Idaho State Department of Education (ID)
Katherine Wiechman	Maize Middle School (KS)

*One panelist did not wish to be listed in the final report.

APPENDIX B

STUDY AGENDA

AGENDA

Praxis® Middle School Mathematics (5164) Standard-Setting Study

Day 1

Welcome and Introduction

Overview of Standard Setting and the *Praxis* Middle School Mathematics Test

Review the *Praxis* Middle School Mathematics Test

Discuss the *Praxis* Middle School Mathematics Test

Lunch

Define the Knowledge/Skills of a Just Qualified Candidate

Break

Define the Just Qualified Candidate *(continued)*

End of Day 1

AGENDA

Praxis® Middle School Mathematics (5164) Standard-Setting Study

Day 2

Overview of Day 2

Define the Just Qualified Candidate *(continued)*

Standard-setting training presentation

Practice Round: Selected-response standard-setting judgments

Break

Practice Round: Data Discussion

Lunch Break

Round 1: Selected-response standard-setting judgments

Break

Round 1: Selected-response standard-setting judgments *(continued)*

End of Day 2

AGENDA

Praxis® Middle School Mathematics (5164) Standard-Setting Study

Day 3

Overview of Day 3

Round 1 Feedback and Round 2 Judgments

Break

Round 1 Feedback and Round 2 Judgments *(continued)*

Break

Feedback on Round 2 Recommended Cut Score

Complete Final Evaluation

End of Study

APPENDIX C

JUST QUALIFIED CANDIDATE DESCRIPTION

Description of the Just Qualified Candidate[10]

A just qualified candidate …

Numbers and Operations

1. Understands proportional reasoning and ratios relationships
2. Understands rational number operations and properties to solve problems (standard and real world)
3. Understand basic concepts of number theory (e.g., prime numbers, factors, exponential rules)
4. Knows how to recognizes the reasonableness of results within the context of a given problem
5. Is familiar with evaluating student work to identify misconceptions and valid explanations of mathematical concepts

Algebra

6. Understands linear equations and systems of two linear equations (solve; represent in multiple forms)
7. Understands representations of one-variable linear inequalities
8. Is familiar with systems of linear inequalities
9. Understands linear relationships in various forms (table, graph, description, equation, etc.)
10. Is familiar with representations of quadratic equations and expressions

Functions

11. Knows how to use and evaluate basic functions that model given information in a variety of contexts through multiple representations (sequence, function notation, tables, and graphs, etc.)
12. Is familiar with absolute value, quadratic, and exponential functions that model given information through multiple representations.
13. Understands the common characteristics and shape of the graph of basic functions, including domain, range, minimum/maximum, slope, and intercepts
14. Is familiar with common characteristics and shape of graph of absolute value, quadratic, and exponential functions

Geometry and Measurement

15. Understands multi-step applications of basic geometric concepts (including area, Pythagorean Theorem, surface area, perimeter, volume, angles/lines, characteristics of shapes, etc.)
16. Knows geometric relationships (e.g., basic transformations, distance, similarity, congruence, systems of measurement)
17. Knows the basic characteristics and properties of circles, as well as triangles, quadrilaterals, and other polygons

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

Description of the Just Qualified Candidate (continued)

A just qualified candidate …

Probability and Statistics

18. Knows how to appropriately collect, interpret, analyze and represent data in various forms and identifies which form is most appropriate in a given situation (e.g., scatter plots, box and whisker plots, stem and leaf, etc.)
19. Understands measures of central tendency
20. Is familiar with variability and can compare two or more data sets
21. Understands how to use basic probability models including those in real world contexts

APPENDIX D

RESULTS

Table D1

Panel Member Demographics (by Panel)

	Panel 1		Panel 2	
	N	**%**	***N***	**%**
Current position				
Teacher	8	53	6	43
Mathematics Instructional Coach	2	13	0	0
Administrator/Department head	1	7	1	7
College faculty	4	27	5	36
Other	0	0	2	14
Race				
White	13	87	11	79
Black or African American	1	7	2	14
American Indian or Alaskan Native	0	0	1	7
Middle Eastern	1	7	0	0
Gender				
Female	13	87	10	71
Male	2	13	4	29
Are you currently certified to teach middle school mathematics in your state?				
Yes	11	73	13	93
No	4	27	1	7
Are you currently teaching middle school mathematics in your state?				
Yes	7	47	7	50
No	8	53	7	50
Are you currently supervising or mentoring other teachers of middle school mathematics?				
Yes	7	47	8	57
No	8	53	6	43
At what K–12 grade level are you currently teaching middle school mathematics?				
Elementary and Middle school	0	0	1	7
Middle school (6–8 or 7–9)	6	40	5	36
Middle and High school	1	7	0	0
High school (9–12 or 10–12)	1	7	0	0
All Grades	1	7	0	0
Other	1	7	0	0
Not currently teaching at the K–12 level	5	33	8	57

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 middle school mathematics?				
3 years or less	1	7	3	21
4–7 years	4	27	0	0
8–11 years	6	40	4	29
12–15 years	2	13	2	14
16 years or more	2	13	5	36
Which best describes the location of your K–12 school?				
Urban	4	27	2	14
Suburban	3	20	2	14
Rural	3	20	2	14
Not currently working at the K–12 level	5	33	8	57
If you are college faculty, are you currently involved in the training/preparation of teacher candidates in middle school mathematic?				
Yes	4	27	4	29
No	0	0	2	14
Not college faculty	11	73	8	57

Table D2

Passing Score Summary by Round of Judgments

	Panel 1		Panel 2	
Panelist	**Round 1**	**Round 2**	**Round 1**	**Round 2**
1	40.15	39.80	30.05	32.70
2	35.90	35.80	38.90	37.20
3	35.20	35.00	37.70	36.50
4	36.45	36.25	33.30	36.00
5	39.20	37.70	37.30	37.90
6	45.55	42.40	40.60	40.45
7	40.00	43.15	36.00	35.95
8	31.85	33.65	38.65	38.25
9	34.60	34.90	31.90	28.10
10	31.20	35.50	38.95	38.50
11	34.05	35.85	34.00	35.00
12	36.65	34.55	36.50	36.40
13	38.80	37.90	36.00	35.50
14	41.40	43.15	38.60	39.40
15	45.35	43.85		
Average	37.76	37.96	36.32	36.28
Lowest	31.20	33.65	30.05	28.10
Highest	45.55	43.85	40.60	40.45
SD	4.31	3.57	3.03	3.06
SEJ	1.11	0.92	0.81	0.82

Table D3

Final Evaluation: Panel 1

	Strongly agree		Agree		Disagree		Strongly disagree	
	N	**%**	***N***	**%**	***N***	**%**	***N***	**%**
• I understood the purpose of this study.	13	87	2	13	0	0	0	0
• The instructions and explanations provided by the facilitators were clear.	9	60	6	40	0	0	0	0
• The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	13	87	2	13	0	0	0	0
• The explanation of how the recommended passing score is computed was clear.	10	67	5	33	0	0	0	0
• The opportunity for feedback and discussion between rounds was helpful.	9	60	6	40	0	0	0	0
• The process of making the standard-setting judgments was easy to follow.	8	53	7	47	0	0	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 just qualified candidate	14	93	1	7	0	0
• The between-round discussions	8	53	7	47	0	0
• The knowledge/skills required to answer each test item	12	80	3	20	0	0
• The passing scores of other panel members	5	33	8	53	2	13
• My own professional experience	14	93	1	7	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?	11	73	4	27	0	0	0	0

	Too low		About right		Too high	
	N	%	*N*	%	*N*	%
• Overall, the recommended passing score is:	2	13	13	87	0	0

Table D4

Final Evaluation: Panel 2

	Strongly agree		Agree		Disagree		Strongly disagree	
	N	**%**	***N***	**%**	***N***	**%**	***N***	**%**
• I understood the purpose of this study.	14	100	0	0	0	0	0	0
• The instructions and explanations provided by the facilitators were clear.	13	93	1	7	0	0	0	0
• The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	13	93	1	7	0	0	0	0
• The explanation of how the recommended passing score is computed was clear.	13	93	1	7	0	0	0	0
• The opportunity for feedback and discussion between rounds was helpful.	14	100	0	0	0	0	0	0
• The process of making the standard-setting judgments was easy to follow.	12	86	2	14	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 just qualified candidate	13	93	1	7	0	0
• The between-round discussions	9	64	5	36	0	0
• The knowledge/skills required to answer each test item	11	79	3	21	0	0
• The passing scores of other panel members	4	29	9	64	1	7
• My own professional experience	11	79	3	21	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?	13	93	1	7	0	0	0	0

	Too low		About right		Too high	
	N	%	*N*	%	*N*	%
• Overall, the recommended passing score is:	0	0	14	100	0	0