

Multistate Standard-Setting Technical Report

***Praxis*® TEACHING READING: K-12 (5206)**

Student and Teacher Assessments: Validity and Test Use

ETS

Princeton, New Jersey

March 2019

EXECUTIVE SUMMARY

To support the decision-making process of education agencies establishing a passing score (cut score) for the *Praxis*[®] Teaching Reading: K-12 (5206) test, research staff from Educational Testing Service (ETS) designed and conducted a multistate standard-setting study.

PARTICIPATING STATES

Panelists from 10 states were recommended by their respective education agencies. The education agencies recommended panelists with (a) experience—as either elementary or secondary reading teachers, reading specialists or as college faculty who prepare reading teachers or specialists—and (b) familiarity with the knowledge and skills required of beginning elementary and secondary reading 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*[®] Teaching Reading: K-12 test, the recommended passing score is 66 out of a possible 107 raw-score points. The scale score associated with a raw score of 66 is 156 on a 100–200 scale.

To support the decision-making process for education agencies establishing a passing score (cut score) for the *Praxis*® Teaching Reading: K-12 (5206) test, research staff from Educational Testing Service (ETS) designed and conducted a multistate standard-setting study in February 2019 in Princeton, New Jersey. Education agencies¹ recommended panelists with (a) experience—as either elementary or secondary reading teachers, reading specialists or as college faculty who prepare reading teachers or specialists—and (b) familiarity with the knowledge and skills required of beginning elementary and secondary reading teachers. Ten states (Table 1) were represented by 17 panelists. (See Appendix A for the names and affiliations of the panelists.)

Table 1
Participating States and Number of Panelists

Arkansas (1 panelist)	Montana (1 panelist)
Iowa (1 panelist)	North Carolina (3 panelists)
Kentucky (2 panelists)	Pennsylvania (2 panelists)
Louisiana (1 panelist)	South Dakota (1 panelist)
Maryland (2 panelists)	West Virginia (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, 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, which represents the combined judgments of a group of experienced educators. Each state may want to consider the recommended passing score but also other sources of information when setting the final *Praxis*® Teaching Reading: K-12 passing score (see Geisinger & McCormick, 2010). A state 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’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

¹ States and jurisdictions that currently use *Praxis* tests were invited to participate in the multistate standard-setting study.

Praxis[®] Teaching Reading: K-12 test score and the latter, the reliability of panelists' passing-score recommendation. The SEM allows a state to recognize that any test score on any standardized test—including a *Praxis*[®] Teaching Reading: K-12 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 to gauge the likelihood that the recommended passing score from the current 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 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 or she should receive a license/certificate, but his or her 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 he or she should not receive a license/certificate, but he or she actually possesses the required knowledge/skills. The state needs to consider which decision error is more important to minimize.

OVERVIEW OF THE *PRAXIS*[®] TEACHING READING: K-12 TEST

The *Praxis*[®] Teaching Reading: K-12 *Study Companion* document (ETS, in press) describes the purpose and structure of the test. In brief, the test focuses on the knowledge and skills a beginning teacher must have to support reading and writing development in elementary and secondary school students.

The 150-minute assessment contains 90 selected-response and 3 constructed-response items² covering six content areas: *Phonological and Phonemic Awareness including Emergent Literacy* (approximately 13 selected-response items), *Phonics and Decoding* (approximately 16 selected-response items), *Fluency and Vocabulary* (approximately 20 selected-response items), *Comprehension of Literacy and Informational Text* (approximately 27 selected-response items), *Writing* (approximately 14 selected-response items) and *Assessment and Instructional Decision Making* (approximately 3 constructed-response items).³ The reporting scale for the *Praxis*[®] Teaching Reading: K-12 test ranges from 100 to 200 points.

PROCESSES AND METHODS

The design of the standard-setting study included an expert panel. 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.

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. ETS content specialists from the assessment development group also provided a brief overview of the test development process. Appendix B shows the agenda for the panel meeting.

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

² Ten of the 90 selected-response items are pretest items and do not contribute to a candidate's score.

³ The number of items for each content area may vary slightly from form to form of the test. Forms of the test may also include an video or audio component.

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.

DESCRIBING 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.

The panel created a description of the just qualified 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 just qualified candidate. The full panel then reconvened and, through whole-group discussion, completed the description of the knowledge and skills 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).

PANELISTS' JUDGMENTS

The *Praxis*[®] Teaching Reading: K-12 test includes both selected-response and constructed-response items (dichotomously- and polytomously-scored, respectively). Panelists received training in two distinct standard-setting approaches: one standard-setting approach for the dichotomously-scored items and another approach for the polytomously-scored items.

A panel's passing score is the sum of the interim passing scores recommended by the panelists for (a) the dichotomously-scored items and (b) the constructed-response items. As with scoring and reporting, the panelists' judgments for the polytomously-scored items were weighted such that they contributed 25% of the overall score.

Dichotomously-scored items. The standard-setting process for the dichotomously-scored items was a probability-based Modified Angoff method (Brandon, 2004; Hambleton & Pitoniak, 2006). In this method, 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 reviewed the description of the just qualified candidate and the item in order to determine the probability that the just qualified candidate would answer the question correctly. To aid the decision-making process, panelists were trained to approach the judgment process in two stages. First, they would consider whether there was a high, moderate, or low chance that the just qualified candidate would correctly answer the question. The following rules of thumb were used to guide their decision:

- If the just qualified candidate would have a low chance of answering correctly, consider the 0 to .30 range of the probability scale.
- If the just qualified candidate would have a moderate chance of answering correctly, consider the .40 to .60 range of the probability scale.
- If the just qualified candidate would have a high chance of answering correctly, consider the .70 to 1 range of the probability scale.

Next, panelists refined their judgment within the range and selected the probability for their judgment. 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, then discussed those judgments and their rationales. The facilitator listened to verify that the panelists followed the training and responded to any questions about how to make standard-setting judgments. Once the practice round was completed, 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.

Polytomously-scored items. An Extended Angoff method (Cizek & Bunch, 2007; Hambleton & Plake, 1995) was used for standard-setting judgments of the constructed-response items. In this method, panelists decide on the assigned score value that would most likely be earned by the just qualified candidate for each constructed-response item. To make their judgments, panelists were trained to review the just qualified candidate description, consider the knowledge and skills required to respond to the constructed-response item, and then consider what is required to earn each point, as described in the rubric. The rubric for a constructed-response item defines (holistically) the quality of the evidence that would merit a response earning a particular score. Each panelist decided on the score most likely to be earned by the just qualified candidate from the possible values a test taker can earn.

A test-taker's response to a constructed-response item is independently scored by two raters, and the sum of the raters' scores is the assigned score⁴; possible scores, therefore, range from zero (both raters assigned a score of zero) to six (both raters assigned a score of three). For their ratings, each panelist decided on the score most likely to be earned by a just qualified candidate from the following possible values: 0, 1, 2, 3, 4, 5, or 6. For each of the constructed-response item, panelists recorded the score (0 through 6) that a just qualified candidate would most likely earn.

After the training, panelists made practice judgments, then discussed those judgments and their rationales. During the practice round, the facilitator listened to verify that the panelists followed the training and responded to any questions about how to make judgments. The practice round concluded with all of the panelists completing 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.

⁴ If the two raters' scores differ by more than one point (non-adjacent), the Chief Reader for that item assigns the score, which is then doubled.

Multiple Rounds. Following the first round of independent judgments (*Round 1*), item-level feedback was provided to the panel. The panelists' judgments were displayed for each item and summarized across panelists. The item-level feedback for the dichotomously-scored items showed the three ranges of judgments (representing the high, moderate, or low chance probabilities for the just qualified candidate to answer correctly) and the number of panelists whose judgments were in those ranges. The average judgment per item was also displayed. To aid the discussion, portions of the item-level feedback were highlighted to illustrate that at least two-thirds of the panel's judgments were in the same probability range. Additionally, panelists were shown which items on the test form did not count towards the total score (their judgments on those items were also not included in the calculation of the recommended score).

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. 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. During the discussion, panelists recorded their Round 2 judgments. They made 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.

RESULTS

EXPERT PANELS

Table 2 presents a summary of the panelists' demographic information. The panel included 17 educators representing 10 states. (See Appendix A for a listing of panelists.) Nine panelists were teachers, seven were college faculty, and one held another position. All of the faculty members' job responsibilities included the training of elementary and secondary reading teachers.

Table 2
Panel Member Demographics

	<i>N</i>	<i>%</i>
Current position		
Teachers	9	53
College faculty	7	41
K-12 Literacy Specialist	1	6
Race		
White or European American	12	71
Black or African American	4	24
American Indian or Alaskan Native	1	6
Gender		
Female	15	88
Male	2	12
Are you currently certified/licensed as a teacher of this subject in your state?		
Yes	16	94
No	1	6
Are you currently teaching this subject in your state?		
Yes	16	94
No	1	6
Are you currently supervising or mentoring other teachers of this subject?		
Yes	15	88
No	2	12
Including this year, how many years of experience do you have teaching this subject?		
3 years or less	0	0
4–7 years	3	18
8–11 years	1	6
12–15 years	5	29
16 years or more	8	47

Table 2 (continued)
Panel Member Demographics

	<i>N</i>	<i>%</i>
For which education level are you currently teaching this subject?		
Elementary (K-5 or K-6)	4	24
Middle School (6-8 or 7-9)	3	18
High School (9-12 or 10-12)	1	6
Middle and High School	1	6
All Grades	1	6
Pre Service Educators	1	6
Not currently working at the K–12 level	6	35
Which best describes the location of your K–12 school?		
Urban	3	18
Suburban	3	18
Rural	5	29
Not currently working at the K–12 level	6	35
If you indicated "College Faculty", are you currently involved in the training or preparation of teacher candidates in this subject?		
Yes	7	41
No	0	0
Not college faculty	10	59

STANDARD-SETTING JUDGMENTS

Table 3 summarizes the standard-setting judgments of panelists. The table shows the passing scores—the number of raw points needed to pass the test—recommended by each panelist.

Table 3 also includes estimate 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.⁵ 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.

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

⁵ 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).

decrease—indicating convergence among the panelists’ judgments—was observed (see Table 3). The Round 2 average score is the panel’s recommended passing score.

Table 3
Passing Score Summary by Round of Judgments

Panelist	Round 1	Round 2
1	66.60	68.00
2	71.98	68.50
3	66.70	68.10
4	63.08	66.07
5	61.47	66.17
6	78.61	76.73
7	74.38	71.50
8	70.03	69.88
9	52.13	55.93
10	54.22	55.52
11	68.60	68.85
12	71.88	71.13
13	72.94	72.16
14	56.57	56.47
15	63.15	64.23
16	55.70	59.77
17	58.82	61.00
Average	65.11	65.88
Lowest	52.13	55.52
Highest	78.61	76.73
SD	7.80	6.21
SEJ	1.89	1.51

The panel’s passing score recommendation for the *Praxis*® Teaching Reading: K-12 test is 65.88 (out of a possible 107 raw-score points). The value was rounded to the next highest whole number, 66, to determine the functional recommended passing score. The scale score associated with 66 raw points is 156.

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 Score⁶***

Recommended passing score (CSEM)		Scale score equivalent
	66 (4.79)	156
-2 CSEM	57	145
-1 CSEM	62	151
+ 1 CSEM	71	162
+ 2 CSEM	76	169

Note. CSEM = conditional standard error(s) of measurement.

FINAL EVALUATIONS

The panelists completed an evaluation at the conclusion of the 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* that they understood the purpose of the study and that the facilitator’s instructions and explanations were clear. All panelists *strongly agreed* that they were prepared to make their standard-setting judgments and that the standard-setting process was easy to follow.

Sixteen of the 17 panelists reported that the description of the just qualified candidate was *very influential* in guiding their standard-setting judgments; the remaining panelist indicated the description was *not influential*. All but one of the panelists reported that between-round discussions were at least *somewhat influential* in guiding their judgments. More than half of the panelists (11 of the 17 panelists) indicated that their own professional experience was *very influential* in guiding their judgments.

Fifteen of the 17 panelists indicated they were *very comfortable* with the passing score they recommended. All but one of the panelists indicated that the recommended passing score was *about right* and the remaining panelist indicated that the passing score was *too high*.

⁶ 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*[®] Teaching Reading: K-12 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*[®] Teaching Reading: K-12 test, the recommended passing score is 66 out of a possible 107 raw-score points. The scale score associated with a raw score of 66 is 156 on a 100–200 scale.

REFERENCES

- Brandon, P. R. (2004). Conclusions about frequently studied modified Angoff standard-setting topics. *Applied Measurement in Education, 17*, 59-88.
- Cizek, G. J., & Bunch, M.B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage.
- ETS. (in press). *The Praxis Series®: The Praxis Study Companion: Teaching Reading: K-12 (5206)*. 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.
- Hambleton, R. K., & Plake, B.S. (1995). Using an extended Angoff procedure to set standards on complex performance assessments. *Applied Measurement in Education, 8*, 41-55.
- 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

<u>Panelist</u>	<u>Affiliation</u>
Latanza Atkins	Southeast Arkansas Educational Cooperative (AR)
Carianne Bernadowski	Robert Morris University (PA)
Lisa Cooper	Louisiana State University Shreveport (LA)
Michelle Devine	North Washington Middle School (KY)
Nicole Finnesand	Tri-Valley School District (SD)
Cequoia Hector	Washington Montessori Elementary School (NC)
Laura Heitritter	Northwestern College (IA)
Helen Hoffner	Holy Family University (PA)
Sean Huneycutt	Wayne County Public Schools (NC)
Mary Lind	Ranson Elementary School (WV)
Daleisha Myers	Prince George's County Public Schools & Bowie State University (MD)
Sarah Pennington	Montana State University (MT)
Toni Poling	Fairmont Senior High School (WV)
Heather Kimberly Dial Sellers	University of North Carolina at Pembroke (NC)
Loray White	Prince George's County Public Schools (MD)
Barbara Wierzbicki	Fairmont State University (WV)
Jarred Winebarger	White Hall Elementary School (KY)

APPENDIX B
STUDY AGENDA

AGENDA

***Praxis*[®] Teaching Reading: K-12 (5206) Standard-Setting Study**

Day 1

Welcome and Introduction

Overview of Standard Setting and the *Praxis*[®] Teaching Reading: K-12 Test

Review and Discussion of the Praxis Teaching Reading: K-12 Test

Break

Define the Knowledge/Skills of a Just Qualified Candidate

Lunch

Define the Knowledge/Skills of a Just Qualified Candidate (continued)

Break

Training and Practice of Modified Angoff Standard-setting Judgments

Round 1 Standard-setting Judgments for the Selected-response Items

Collect Materials; End of Day 1

AGENDA

***Praxis*[®] Teaching Reading: K-12 (5206) Standard-Setting Study**

Day 2

Overview of Day 2

Training and Practice in Extended Angoff Standard-setting Judgments

Round 1 Standard Setting Judgments for Constructed-Response Items

Break

Round 1 Feedback & Round 2 Judgments

Lunch

Feedback on Round 2 Recommended Cut Score

Complete Final Evaluation

Collect Materials; End of Study

APPENDIX C

JUST QUALIFIED CANDIDATE DESCRIPTION

Description of the Just Qualified Candidate⁷

A just qualified candidate...

1. Knows how to use and interpret literacy assessment practices to inform instruction.
2. Knows some methods and strategies for all five essential components of effective reading instruction (digital & print) across grade levels (phonemic awareness, phonics, fluency, vocabulary, and comprehension).
3. Knows how to differentiate to meet the needs of diverse learners.
4. Is familiar with how to teach writing as a recursive process.
5. Is familiar with the integration of reading, writing, speaking and listening skills as being integral to literacy instruction.
6. Knows some instructional methods for supporting reading and writing in varied contexts and disciplinary domains.

⁷ Description of the just qualified candidate focuses on the knowledge/skills that differentiate a *just* from a *not quite* qualified candidate.

APPENDIX D

FINAL EVALUATION RESULTS

Table D1***Final Evaluation***

	Strongly agree		Agree		Disagree		Strongly disagree	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
• I understood the purpose of this study.	17	100	0	0	0	0	0	0
• The instructions and explanations provided by the facilitators were clear.	17	100	0	0	0	0	0	0
• The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	17	100	0	0	0	0	0	0
• The explanation of how the recommended passing score is computed was clear.	17	100	0	0	0	0	0	0
• The opportunity for feedback and discussion between rounds was helpful.	16	94	1	6	0	0	0	0
• The process of making the standard-setting judgments was easy to follow.	17	100	0	0	0	0	0	0
• I understood how to use the survey software	17	100	0	0	0	0	0	0

Table D1 (continued)

Final Evaluation

How influential was each of the following factors in guiding your standard-setting judgments?	Very influential		Somewhat influential		Not influential			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• The description of the just qualified candidate	16	94	0	0	1	6		
• The between-round discussions	10	59	6	35	1	6		
• The knowledge/skills required to answer each test item	15	88	1	6	1	6		
• The passing scores of other panel members ⁸	4	24	9	53	3	18		
• My own professional experience	11	65	5	29	1	6		
	Very comfortable		Somewhat comfortable		Somewhat uncomfortable		Very uncomfortable	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
• Overall, how comfortable are you with the panel's recommended passing score?	15	88	2	12	0	0	0	0
	Too low		About right		Too high			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• Overall, the recommended passing score is:	0	0	16	94	1	6		

⁸ One of the panelists did not answer this question