Listening. Learning. Leading.®

Multistate Standard-Setting Technical Report

PRAXIS® COMPUTER SCIENCE (5652)

Educational Testing Service

Princeton, New Jersey

February 2018

EXECUTIVE SUMMARY

To support the decision-making process of education agencies establishing a passing score (cut score) for the *Praxis*® Computer Science (5652) test, research staff from Educational Testing Service (ETS) designed and conducted a multistate standard-setting study.

PARTICIPATING STATES

Panelists from 17 states and Washington, DC were recommended by their respective education agencies. The education agencies recommended panelists with (a) experience as either computer science teachers or college faculty who prepare computer science teachers and (b) familiarity with the knowledge and skills required of beginning computer science 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* Computer Science test, the recommended passing score[1] is 47 out of a possible 80 raw-score points. The scale score associated with a raw score of 47 is 149 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*® Computer Science (5652) test, research staff from ETS designed and conducted a multistate standard-setting study in January 2018 in Princeton, New Jersey. Education agencies[2] recommended panelists with (a) experience as either computer science teachers or college faculty who prepare computer science teachers and (b) familiarity with the knowledge and skills required of beginning computer science teachers. Seventeen states and Washington, DC (Table 1) were represented by 36 panelists. (See Appendix A for the names and affiliations of the panelists.)

Table 1
Participating Jurisdictions and Number of Panelists

Alabama (2 panelists)	Pennsylvania (3 panelists)
Arkansas (2 panelists)	South Carolina (1 panelist)
Georgia (4 panelists)	South Dakota (1 panelist)
Idaho (2 panelists)	Tennessee (2 panelists)
Kentucky (3 panelists)	Utah (2 panelists)
Maryland (2 panelists)	Virginia (2 panelists)
Nevada (1 panelist)	Washington, DC (2 panelists)
New Jersey (2 panelists)	West Virginia (2 panelists)
North Dakota (1 panelist)	Wisconsin (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 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

[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 recommened passing scores for each panel are presented.

Praxis Computer Science 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* Computer Science 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* Computer Science 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 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 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 jurisdiction needs to consider which decision error is more important to minimize.

OVERVIEW OF THE *PRAXIS* COMPUTER SCIENCE TEST

The *Praxis* Study Companion for the Computer Science (5652) test (ETS, in press) describes the purpose and structure of the test. In brief, the test is designed to assess the computer science knowledge and competencies necessary for a beginning teacher of secondary school computer science.

The three-hour assessment contains 100 selected-response items[4] covering five content areas: *Impacts of Computing* (approximately 15 items), *Algorithms and Computational Thinking* (approximately 25 items), *Programming* (approximately 30 items), *Data* (approximately 15 items), and *Computing Systems and Networks* (approximately 15 items).[5] The reporting scale for the *Praxis* Computer Science test ranges from 100 to 200 scale-score points.

PROCESSES AND METHODS

The design of the standard-setting study included two 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.

The standard-setting study began with a welcome and introduction by the meeting facilitators. The facilitators 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.

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.

[4] Twenty of the 100 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.

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 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 a description of the just qualified candidate — the knowledge/skills that differentiate a *just* from a *not quite* qualified candidate. To create this description, they first split into smaller groups to consider the just qualified candidate. Then they reconvened and, through whole-group discussion, created the description of the just qualified candidate to use for the remainder of the study. After the description was completed, panelists were split into two, distinct panels that worked separately 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 standard-setting process for the *Praxis* Computer Science 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. Then the panelists estimated what chance a just qualified candidate would have of answering 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 evaulation 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 36 educators representing 17 states and Washington, DC. (See Appendix A for a listing of panelists.) Twenty-two panelists were teachers, one was an administrator or department head, nine were college faculty, and four held another position. All of the faculty members' job responsibilities included the training of computer science teachers.

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

Table 2

Panel Member Demographics (Across Panels)

	N	%
Current position		
Teacher	22	61
Administrator/Department Head	1	3
College Faculty	9	25
Other	4	11
Race		
White	24	67
Black or African American	4	11
Hispanic or Latino	1	3
Asian or Asian American	5	14
Other	1	3
No Response	1	3
Gender		
Female	18	50
Male	18	50
Are you currently certified to teach this subject in your state?		
Yes	20	56
No	16	44
Are you currently teaching this subject in your state?		
Yes	32	89
No	4	11
Are you currently supervising or mentoring other teachers of this subject?		
Yes	20	56
No	16	44
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	56
Middle and High School	1	3
All Grades	1	3
Other	3	8
Not currently teaching at the K–12 level	10	28

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	7	19
4–7 years	9	25
8–11 years	7	19
12–15 years	5	14
16 years or more	8	22
Which best describes the location of your K–12 school?		
Urban	7	19
Suburban	12	33
Rural	8	22
Not currently working at the K–12 level	9	25
If you are college faculty, are you currently involved in the training/preparation of teacher candidates in this subject?		
Yes	7	19
No	2	6
Not college faculty	27	75

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 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).

[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	44.48	48.72
Lowest	35.70	39.90
Highest	54.00	55.65
SD	5.65	4.38
SEJ	1.33	1.03

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 each panel (see Table D2 in Appendix D). The Round 2 average score is the panel's recommended passing score.

The panels' passing score recommendations for the *Praxis* Computer Science test are 44.48 for Panel 1 and 48.72 for Panel 2 (out of a possible 80 raw-score points). The values were rounded to the next highest whole number, to determine the functional recommended passing score — 45 for Panel 1 and 49 for Panel 2. The scale scores associated with 45 and 49 raw points are 145 and 152, 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* Computer Science test is 46.60 (out of a possible 80 raw-score points). The value was rounded to 47 (next highest raw score) to determine the functional recommended passing score. The scale score associated with 47 raw points is 149.

Table 4 presents the estimated conditional standard error of measurement (CSEM) around the recommended passing score (the average across the two panels) 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[7]

Recommended passing score (CSEM)		Scale score equivalent
47 (4.43)		149
-2 CSEM	39	135
-1 CSEM	43	142
+ 1 CSEM	52	158
+ 2 CSEM	56	165

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 their 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 36 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 (21 of the 36 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 scale scores.

All but two of the panelists, both on Panel 1, indicated they were at least *somewhat comfortable* with the passing score they recommended; 23 of the 36 panelists were *very comfortable*. Thirty-two of the 36 panelists indicated the recommended passing score was *about right*; four panelists one Panel 1 indicated that the passing score was *too low*.

SUMMARY

To support the decision-making process for education agencies establishing a passing score (cut score) for the *Praxis* Computer Science 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* Computer Science test, the recommended passing score[8] is 47 out of a possible 80 raw-score points. The scale score associated with a raw score of 47 is 149 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®: Study Companion: Computer Science (5652)*. 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**
Jason Beach	Tennessee Tech University (TN)
Patricia Beach	Georgia Department of Education (GA)
Nanette Brothers	Sandpoint High School (ID)
Kent Brown	New Rockford - Sheyenne School District 2 (ND)
Cindi Chang	Nevada Department of Education (NV)
Drew Fulkerson	Bowling Green High School (KY)
Mark Grammer	Uintah High School (UT)
Rabiah Harris	Dunbar High School/District of Columbia Public Schools (DC)
Lila Holt	University of Tennessee (TN)
Robert Honomichl	Dakota State University (SD)
Jennifer Howard	West Jessamine Middle School (KY)
Lori Hunt	Middleton High School (WI)
Amal Ileiwat	Paterson Public Schools (NJ)
Amit Jain	Boise State University (ID)
Russel Johnson	Auburn High School (AL)
Robert Juranitch	University School of Milwaukee (WI)
Lisa Kovalchick	California University of Pennsylvania (PA)
Yesem Kurt Peker	Columbus State University (GA)
Yu Liu	Fayette County Board of Education (GA)
Curt Minich	Wyomissing Area High School (PA)
Jigish Patel	Northwest Arkansas Education Service Cooperative (AR)

Participating Panelists With Affiliation (continued)

Panelist	**Affiliation**
Jandelyn (Jan) Plane	University of Maryland College Park (MD)
Douglas Poland	Stone Bridge High School (VA)
Lauren Poutasse	Delaware County Intermediate Unit (PA)
Cong Pu	Marshall University (WV)
Nicole Reitz-Larsen	West High School (UT)
Andrea Robertson	Wheaton High School (MD)
Justin Smith	Metcalfe County High School (KY)
Kyle Tower	Lee-Davis High School (VA)
Donnita Tucker	Francis Marion School (AL)
Blake Vaught	Academy for the Arts, Science, and Technology (SC)
Kelly L. Vostal	West Windsor-Plainsboro Board of Education (NJ)
Paulus Wahjudi	Marshall University (WV)
Karl Walker	University of Arkansas at Pine Bluff (AR)
Shirl Williams	Houston County High School (GA)
Melanie Wiscount	District of Columbia Public Schools (DC)

APPENDIX B

STUDY AGENDA

AGENDA

Praxis® Computer Science (5652)
Standard-Setting Study

Day 1

Welcome and Introduction

Overview of Standard Setting and the *Praxis* Computer Science Test

Review the *Praxis* Computer Science Test

Discuss the *Praxis* Computer Science Test

Define the Knowledge/Skills of a Just Qualified Candidate

Standard-Setting Training

Round 1 Standard Setting Judgments

Collect Materials; End of Day 1

Day 2

Overview of Day 2

Round 1 Feedback and Round 2 Judgments

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 Candidate[9]

A just qualified candidate …

I. Impacts of Computing

1. Is familiar with harmful and beneficial impacts of contemporary computing on society, economy, and culture
2. Knows challenges to equal access to computing among different groups and impacts of those obstacles and familiar with existing strategies to address them
3. Is familiar with basic issues regarding intellectual property and ethics in computing
4. Knows basic trade-offs involved in privacy and security issues regarding the acquisition, use and disclosure of information in a digital world

II. Algorithms

1. Knows how to use pattern recognition, problem decomposition and abstraction
2. Is familiar with how to analyze algorithms expressed in multiple formats (natural language, flowcharts, pseudocode)
3. Is familiar with basic algorithms (e.g., count, sum, swap, search, sort)

III. Programming

1. Understands the three basic constructs used in programming: sequence, selection, and iteration
2. Understands how to use variables, a variety of data types, and the basic array/list data structure
3. Knows how to implement, debug, trace and test computer programs for correctness
4. Knows how to write and call procedures with parameters and return values

IV. Data

1. Knows how data is represented by computers
2. Is familiar with how computers are used to transform (e.g., number conversion, binary, encryption) and process data
3. Is familiar with the applications of computing in modeling and simulation

V. Computing Systems and Networks

1. Knows the basic hardware and software components of a computer and their functions
2. Is familiar with networking, including security issues and the Internet

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

APPENDIX D

RESULTS

Table D1

Panel Member Demographics (by Panel)

	Panel 1		Panel 2	
	N	**%**	***N***	**%**
Current position				
Teacher	12	67	10	56
Administrator/Department Head	0	0	1	6
College Faculty	4	22	5	28
Other	2	11	2	11
Race				
White	11	61	13	72
Black or African American	2	11	2	11
Hispanic or Latino	1	6	0	0
Asian or Asian American	3	17	2	11
No Response	1	6	0	0
Other	0	0	1	6
Gender				
Female	9	50	9	50
Male	9	50	9	50
Are you currently certified to teach this subject in your state?				
Yes	11	61	9	50
No	7	39	9	50
Are you currently teaching this subject in your state?				
Yes	15	83	17	94
No	3	17	1	6
Are you currently supervising or mentoring other teachers of this subject?				
Yes	10	56	10	56
No	8	44	8	44
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)	11	61	9	50
Middle and High School	0	0	1	6
All Grades	0	0	1	6
Other	1	6	2	11
Not currently teaching at the K–12 level	5	28	5	28

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	5	28	2	11
4–7 years	5	28	4	22
8–11 years	3	17	4	22
12–15 years	3	17	2	11
16 years or more	2	11	6	33
Which best describes the location of your K–12 school?				
Urban	4	22	3	17
Suburban	7	39	5	28
Rural	3	17	5	28
Not currently working at the K–12 level	4	22	5	28
If you are college faculty, are you currently involved in the training/preparation of teacher candidates in this subject?				
Yes	2	11	5	28
No	2	11	0	0
Not college faculty	14	78	13	72

Table D2

Passing Score Summary by Round of Judgments

	Panel 1		Panel 2	
Panelist	**Round 1**	**Round 2**	**Round 1**	**Round 2**
1	44.40	42.40	49.25	48.85
2	35.65	35.70	55.50	52.40
3	35.25	37.15	51.35	54.40
4	39.10	38.80	45.45	46.35
5	37.45	35.95	51.35	51.65
6	36.65	39.45	43.50	44.10
7	47.05	49.30	58.10	55.65
8	54.70	54.00	38.20	45.65
9	43.40	45.50	54.40	51.40
10	56.65	53.85	54.50	54.60
11	44.50	43.00	58.20	52.75
12	44.35	47.35	50.25	48.85
13	46.00	45.50	45.70	45.35
14	50.70	50.30	46.60	47.70
15	47.65	46.85	35.90	39.90
16	44.15	48.90	45.70	46.30
17	42.25	42.55	47.90	48.00
18	40.00	44.10	43.90	43.00
Average	43.88	44.48	48.65	48.72
Lowest	35.25	35.70	35.90	39.90
Highest	56.65	54.00	58.20	55.65
SD	6.10	5.65	6.26	4.38
SEJ	1.44	1.33	1.47	1.03

Table D3

Final Evaluation: Panel 1

	Strongly agree		Agree		Disagree		Strongly disagree	
	N	**%**	***N***	**%**	***N***	**%**	***N***	**%**
• I understood the purpose of this study.	14	78	4	22	0	0	0	0
• The instructions and explanations provided by the facilitators were clear.	16	89	2	11	0	0	0	0
• The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	12	67	6	33	0	0	0	0
• The explanation of how the recommended passing score is computed was clear.	12	67	6	33	0	0	0	0
• The opportunity for feedback and discussion between rounds was helpful.	15	83	3	17	0	0	0	0
• The process of making the standard-setting judgments was easy to follow.	13	72	5	28	0	0	0	0
• I understood how to use the survey software.	16	89	2	11	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	10	56	8	44	0	0
• The between-round discussions	8	44	10	56	0	0
• The knowledge/skills required to answer each test item	14	78	4	22	0	0
• The passing scores of other panel members	2	11	13	72	3	17
• My own professional experience	12	67	6	33	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?	9	50	7	39	2	11	0	0

	Too low		About right		Too high	
	N	%	*N*	%	*N*	%
• Overall, the recommended passing score is:	4	22	14	78	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.	18	100	0	0	0	0	0	0
• The instructions and explanations provided by the facilitators were clear.	18	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.	15	83	3	17	0	0	0	0
• The explanation of how the recommended passing score is computed was clear.	16	89	2	11	0	0	0	0
• The opportunity for feedback and discussion between rounds was helpful.	17	94	1	6	0	0	0	0
• The process of making the standard-set`ting judgments was easy to follow.	15	83	3	17	0	0	0	0
• I understood how to use the survey software.	17	94	1	6	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	17	94	1	6	0	0
• The between-round discussions	13	72	4	22	1	6
• The knowledge/skills required to answer each test item	14	78	4	22	0	0
• The passing scores of other panel members	3	17	14	78	1	6
• 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?	14	78	4	22	0	0	0	0

	Too low		About right		Too high	
	N	%	*N*	%	*N*	%
• Overall, the recommended passing score is:	0	0	18	100	0	0