



Listening. Learning. Leading.

Multistate Standard-Setting Technical Report

PRAXIS™ FAMILY AND CONSUMER SCIENCES (5122)

Licensure and Credentialing Research

ETS

Princeton, New Jersey

October 2013

EXECUTIVE SUMMARY

To support the decision-making process of education agencies establishing a passing score (cut score) for the Praxis™ Family and Consumer Sciences (5122) test, research staff from Educational Testing Service (ETS) designed and conducted a multistate standard-setting study.

PARTICIPATING STATES

Panelists from 20 states were recommended by their respective education agencies. The education agencies recommended panelists with (a) experience as either Family and Consumer Sciences teachers or college faculty who prepare Family and Consumer Sciences teachers and (b) familiarity with the knowledge and skills required of beginning Family and Consumer Sciences 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 Family and Consumer Sciences test, the recommended passing score¹ is 67 out of a possible 110 raw-score points. The scaled score associated with a raw score of 67 is 153 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™ Family and Consumer Sciences (5122) test, research staff from ETS designed and conducted a multistate standard-setting study in September 2013 in Princeton, New Jersey. Education agencies² recommended panelists with (a) experience as either Family and Consumer Sciences teachers or college faculty who prepare Family and Consumer Sciences teachers and (b) familiarity with the knowledge and skills required of beginning Family and Consumer Sciences teachers. Twenty states (Table 1) were represented by 32 panelists. (See Appendix A for the names and affiliations of the panelists.)

Table 1
Participating States and Number of Panelists

Arkansas (2 panelists)	North Dakota (2 panelists)
Connecticut (2 panelists)	Nebraska (2 panelists)
Delaware (1 panelist)	Nevada (1 panelist)
Idaho (1 panelist)	Pennsylvania (1 panelist)
Kansas (2 panelists)	Rhode Island (2 panelists)
Kentucky (2 panelists)	South Dakota (1 panelist)
Louisiana (2 panelists)	Utah (2 panelists)
Maryland (1 panelist)	Virginia (2 panelists)
Maine (2 panelists)	Wisconsin (1 panelist)
North Carolina (2 panelists)	Wyoming (1 panelist)

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,³ which represents the combined judgments of two panels of experienced educators. Each state may want to consider the

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

³ In addition to the recommended passing score averaged across the two panels, the recommended passing scores for each panel are presented.

recommended passing score but also other sources of information when setting the final Praxis Family and Consumer Sciences 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 Praxis Family and Consumer Sciences 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 Family and Consumer Sciences 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 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 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 needs to consider which decision error is more important to minimize.

OVERVIEW OF THE PRAXIS FAMILY AND CONSUMER SCIENCES TEST

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

The two-hour assessment contains 120 selected-response items⁴ covering six content areas: *Food and Nutrition* (approximately 26 items), *Housing and Interior Design* (approximately 14 items), *Textiles, Fashion, and Apparel* (approximately 14 items), *Human Development and Interpersonal Relationships* (approximately 24 items), *Foundations of Family and Consumer Sciences Education* (approximately 20 items), and *Resource Management* (approximately 22 items).⁵ The reporting scale for the Praxis Family and Consumer Sciences test ranges from 100 to 200 scaled-score points.

PROCESSES AND METHODS

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

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

REVIEWING THE TEST

The standard-setting panelists first reviewed 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 Ten of the 120 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 TARGET CANDIDATE

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

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

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

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

PANELISTS' JUDGMENTS

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

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

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

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

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

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

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

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

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

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

RESULTS

EXPERT PANELS

Table 2 presents a summary of the panelists' demographic information. The panel included 32 educators representing 20 states. (See Appendix A for a listing of panelists.) Twenty-two panelists were teachers, seven were college faculty, and three were administrators or department heads. All seven faculty members' job responsibilities included the training of Family and Consumer Sciences 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)

	<i>N</i>	<i>%</i>
Current position		
Teacher	22	69%
Administrator/Department Head	3	9%
College Faculty	7	22%
Race		
White	29	91%
Black	3	9%
Gender		
Female	31	97%
Male	1	3%
Are you currently certified to teach this subject in your state?		
Yes	27	84%
No	5	16%
Are you currently teaching this subject in your state?		
Yes	30	94%
No	2	6%
Are you currently supervising or mentoring other teachers of this subject?		
Yes	18	56%
No	14	44%
At what K–12 grade level are you currently teaching this subject?		
Middle School (6–8 or 7–9)	5	16%
High School (9–12 or 10–12)	16	50%
Middle and High School	2	6%
Other	1	3%
Not currently teaching at the K–12 level	8	25%

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

	<i>N</i>	<i>%</i>
Including this year, how many years of experience do you have teaching this subject?		
3 years or less	3	9%
4–7 years	6	19%
8–11 years	7	22%
12–15 years	4	13%
16 years or more	12	38%
Which best describes the location of your school/institution?		
Urban	9	28%
Suburban	13	41%
Rural	10	31%
If you are college faculty, are you currently involved in the training/preparation of teacher candidates in this subject?		
Yes	7	22%
No	0	0%
Not college faculty	25	78%

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	65.87	67.32
Lowest	51.95	56.00
Highest	79.15	79.30
SD	6.58	5.52
SEJ	1.59	1.42

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 Family and Consumer Sciences test are 65.87 for Panel 1 and 67.32 for Panel 2 (out of a possible 110 raw-score points). The values were rounded to the next highest whole number, to determine the functional recommended passing score — 66 for Panel 1 and 68 for Panel 2. The scaled scores associated with 66 and 68 raw points are 151 and 154, 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 Family and Consumer Sciences test is 66.59 (out of a possible 110 raw-score points). The value was rounded to 67 (next highest raw score) to determine the functional recommended passing score. The scaled score associated with 67 raw points is 153.

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

Table 4***Passing Scores Within 1 and 2 CSEMs of the Recommended Passing Score⁷***

Recommended passing score (CSEM)	Scale score equivalent
67 (5.14)	153
-2 CSEMs	140
-1 CSEM	146
+ 1 CSEM	160
+ 2 CSEMs	166

Note. CSEM = conditional standard error of measurement.

FINAL EVALUATIONS

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

Panelists were also shown the panel's recommended passing score and asked (a) how comfortable they are with the recommended passing score and (b) if they think the score was too high, too low, or about right. A summary of the final evaluation results is presented in Appendix D (Tables D3 and D4).

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

Twenty-nine of the 32 panelists indicated they were at least *somewhat comfortable* with the passing score they recommended; 19 of the 32 panelists were *very comfortable*. Twenty-eight of the 32 panelists indicated the recommended passing score was *about right* with the remaining four panelists indicated that the passing score was *too low*.

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

SUMMARY

To support the decision-making process for education agencies establishing a passing score (cut score) for the Praxis Family and Consumer Sciences 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 Family and Consumer Sciences test, the recommended passing score⁸ is 67 out of a possible 110 raw-score points. The scaled score associated with a raw score of 67 is 153 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: Family and Consumer Sciences (5122)*. 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

<u>Panelist</u>	<u>Affiliation</u>
Danielle Albanese	Washington Middle School/Meriden BOE (CT)
Karma Bateman	Jordan High School (UT)
Mari Borr	North Dakota State University (ND)
Kristen Burnett	Laramie County School District 1 Central High School (WY)
Kathy Croxall	Western Kentucky University (KY)
Catrina Cheek-Jones	Richmond Public Schools (VA)
Annette Dutton	Century High School (ID)
Carol Erwin	Wayne State College (NE)
Derrick L. Fogg	Warren County School (NC)
Lindsay Frisbie	South Middleton School District (PA)
Kathleen Hudson	Coventry High School (RI)
Beth Ireland	Kansas State University- Research & Extension (KS)
Maria Keels	Cumberland County Schools/MacWilliams Middle School (NC)
Tracy Kern	Harrisburg High School (SD)
Brenda A. Martin	University of Arkansas at Pine Bluff (AR)
Ronda Matthews	Appomattox County High School (VA)
Yvonne Moody	Chadron State College (NE)
Katrina Moore	Bossier High School (LA)
Christine Moore	Brigham Young University (UT)
Kathy Morgan	Carroll County Public School (MD)
Marlene Prichard	Barrington Middle School (RI)
Sarah Raikes	Washington County High School (KY)
Magan Randall	Fayetteville Public Schools (AR)
Anne Scharmberg	Brandywine High School (DE)
Carla Selberg	Brunswick High School (ME)
Stephanie Sevigny	Velva High School (ND)
Lola Shipp	Liberty High School, Clark County (NV)
Lorraine Tanguay	Mountain Valley High School (ME)
Susan Turgeson	University of WI-Stevens Point (WI)
Janice M. Uerz	Manchester High School (CT)
Amy Weems	Ouachita Junior High School and University of Louisiana at Monroe (LA)
Sally Yahnke	Kansas State University (KS)

APPENDIX B

STUDY AGENDA

AGENDA

Praxis Family and Consumer Sciences (5122) Standard-Setting Study

Day 1

Welcome and Introduction

Overview of Standard Setting and the Praxis Family and Consumer Sciences Test

Review the Praxis Family and Consumer Sciences Test

Discuss the Praxis Family and Consumer Sciences Test

Define the Knowledge/Skills of a Target Candidate

Lunch

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

Break

Standard-Setting Training

Round 1 Standard Setting Judgments

Collect Materials; End of Day 1

AGENDA

Praxis Family and Consumer Sciences (5122) Standard-Setting Study

Day 2

Overview of Day 2

Round 1 Feedback and Round 2 Judgments

Break

Round 1 Feedback and Round 2 Judgments (continued)

Lunch

Feedback on Round 2 Recommended Passing Score

Complete Final Evaluation

Collect Materials; End of Study

TARGET CANDIDATE DESCRIPTION

Description of the Target Candidate⁹

A target candidate ...

- I Food & Nutrition
 - A. Culinary Arts & Food Service
 - 1. Understands basic safety and sanitation procedures
 - 2. Knows general concepts of food science
 - 3. Understands food preparation techniques, knowledge, & skills
 - 4. Knows basic process skills needed in the hospitality industry
 - B. Nutrition and Wellness
 - 1. Understands basic nutrients and dietary guidelines recommended throughout the life cycle
- II Housing & Interior Design
 - 1. Knows the various factors that affect housing choices through the lifespan
 - 2. Understands the elements and principles of design
- III Textiles, Fashion, and Apparel
 - 1. Knows the basic construction techniques, care, and design of textile products and apparel
- IV Human Development and Interpersonal Relationships
 - A. Families and Relationships
 - 1. Knows functions, strategies and resources for promoting relationships and communication skills for various types of families and individuals.
 - B. Human Development and Early Childhood Education
 - 1. Understands the foundational theories related to the stages, characteristics, and interrelatedness of physical, social, emotional, moral, and cognitive development throughout the life cycle
 - 2. Is familiar with the components of a curriculum that promote and address the diverse needs of children
- V. Foundations of Family and Consumer Sciences Education
 - 1. Knows how to implement core FACS academic standards into authentic, student centered learning experiences and assessments
 - 2. Is familiar with career paths and organizations within FACS disciplines from current and historical perspectives.
- VI. Resource Management
 - A. Financial Literacy
 - 1. Understands essential financial literacy concepts and their impacts on financial management throughout the life cycle
 - B. Careers and Consumer Education
 - 1. Knows consumer rights, roles and responsibilities related to individual, family and government resources
 - 2. Knows decision-making and problem-solving processes for family, career and consumer issues

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

APPENDIX D

RESULTS

Table D1
Panel Member Demographics (by Panel)

	Panel 1		Panel 2	
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
Current position				
Teacher	12	71%	10	67%
Administrator/Department Head	2	12%	1	7%
College Faculty	3	18%	4	27%
Race				
White	15	88%	14	93%
Black	2	12%	1	7%
Gender				
Female	17	100%	14	93%
Male	0	0%	1	7%
Are you currently certified to teach this subject in your state?				
Yes	15	88%	12	80%
No	2	12%	3	20%
Are you currently teaching this subject in your state?				
Yes	17	100%	13	87%
No	0	0%	2	13%
Are you currently supervising or mentoring other teachers of this subject?				
Yes	9	53%	9	60%
No	8	47%	6	40%
At what K–12 grade level are you currently teaching this subject?				
Middle School (6–8 or 7–9)	3	18%	2	13%
High School (9–12 or 10–12)	9	53%	7	47%
Middle and High School	1	6%	1	7%
Other	0	0%	1	7%
Not currently teaching at the K–12 level	4	24%	4	27%

Table D1 (continued)***Panel Member Demographics (by Panel)***

	Panel 1		Panel 2	
	<i>N</i>	%	<i>N</i>	%
Including this year, how many years of experience do you have teaching this subject?				
3 years or less	3	18%	0	0%
4–7 years	3	18%	3	20%
8–11 years	4	24%	3	20%
12–15 years	2	12%	2	13%
16 years or more	5	29%	7	47%
Which best describes the location of your school/institution?				
Urban	5	29%	4	27%
Suburban	7	41%	6	40%
Rural	5	29%	5	33%
If you are college faculty, are you currently involved in the training/preparation of teacher candidates in this subject?				
Yes	3	18%	4	27%
No	0	0%	0	0%
Not college faculty	14	82%	11	73%

Table D2***Passing Score Summary by Round of Judgments***

Panelist	Panel 1		Panel 2	
	Round 1	Round 2	Round 1	Round 2
1	59.75	60.80	57.15	66.60
2	84.65	79.15	70.30	70.40
3	73.60	74.10	65.40	65.50
4	51.70	54.60	67.20	67.90
5	65.90	65.40	69.15	68.05
6	69.75	69.00	69.05	69.05
7	69.60	70.50	62.10	62.60
8	62.70	62.10	74.20	76.40
9	48.45	51.95	51.25	56.00
10	50.20	62.20	58.40	63.00
11	67.00	66.50	61.00	65.30
12	65.30	65.00	63.80	64.20
13	73.90	71.05	66.30	67.80
14	67.75	67.60	81.65	79.30
15	63.85	65.65	65.70	67.70
16	65.10	65.10		
17	70.80	69.10		
Average	65.29	65.87	65.51	67.32
Lowest	48.45	51.95	51.25	56.00
Highest	84.65	79.15	81.65	79.30
SD	9.12	6.58	7.30	5.52
SEJ	2.21	1.59	1.89	1.42

Table D3***Final Evaluation: Panel 1***

	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 facilitator were clear.	15	88%	2	12%	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	88%	2	12%	0	0%	0	0%
• The explanation of how the recommended passing score is computed was clear.	16	94%	1	6%	0	0%	0	0%
• The opportunity for feedback and discussion between rounds was helpful.	17	100%	0	0%	0	0%	0	0%
• The process of making the standard-setting judgments was easy to follow.	10	59%	7	41%	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			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• The description of the target candidate	16	94%	1	6%	0	0%		
• The between-round discussions	7	41%	10	59%	0	0%		
• The knowledge/skills required to answer each test item	13	76%	4	24%	0	0%		
• The passing scores of other panel members	4	24%	13	76%	0	0%		
• My own professional experience	13	76%	3	18%	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?	10	59%	4	24%	2	12%	1	6%
	Too low		About right		Too high			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• Overall, the recommended passing score is:	4	24%	13	76%	0	0%		

Table D4***Final Evaluation: Panel 2***

	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.	15	100%	0	0%	0	0%	0	0%
• The instructions and explanations provided by the facilitator were clear.	15	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.	14	93%	1	7%	0	0%	0	0%
• The explanation of how the recommended passing score is computed was clear.	14	93%	1	7%	0	0%	0	0%
• The opportunity for feedback and discussion between rounds was helpful.	15	100%	0	0%	0	0%	0	0%
• The process of making the standard-setting judgments was easy to follow.	12	80%	3	20%	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			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• The description of the target candidate	15	100%	0	0%	0	0%		
• The between-round discussions	11	73%	4	27%	0	0%		
• The knowledge/skills required to answer each test item	14	93%	1	7%	0	0%		
• The passing scores of other panel members	3	20%	10	67%	2	13%		
• My own professional experience	5	33%	10	67%	0	0%		
	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?	9	60%	6	40%	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%	15	100%	0	0%		