Listening. Learning. Leading.

Multistate Standard-Setting Technical Report

PRAXIS™ AGRICULTURE (5701)

Licensure and Credentialing Research

ETS

Princeton, New Jersey

November 2013

EXECUTIVE SUMMARY

To support the decision-making process of education agencies establishing a passing score (cut score) for the Praxis™ Agriculture (5701) test, research staff from Educational Testing Service (ETS) designed and conducted a multistate standard-setting study.

PARTICIPATING STATES

Panelists from 18 states were recommended by their respective education agencies. The education agencies recommended panelists with (a) experience as either agriculture teachers or college faculty who prepare agriculture teachers and (b) familiarity with the knowledge and skills required of beginning agriculture 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 Agriculture test, the recommended passing score[1] is 64 out of a possible 110 raw-score points. The scaled score associated with a raw score of 64 is 147 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™ Agriculture (5701) test, research staff from ETS designed and conducted a multistate standard-setting study in October 2013 in Princeton, New Jersey. Education agencies[2] recommended panelists with (a) experience as either agriculture teachers or college faculty who prepare agriculture teachers and (b) familiarity with the knowledge and skills required of beginning agriculture teachers. Eighteen states (Table 1) were represented by 28 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)
Delaware (2 panelists)	Pennsylvania (2 panelists)
Iowa (1 panelist)	South Carolina (1 panelist)
Kansas (1 panelist)	South Dakota (2 panelists)
Kentucky (2 panelists)	Tennessee (2 panelists)
Louisiana (1 panelist)	Utah (2 panelists)
Maryland (1 panelist)	West Virginia (1 panelist)
Nebraska (2 panelists)	Wisconsin (2 panelists)
Nevada (1 panelist)	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 recommended passing score but also other sources of information when setting the final Praxis Agriculture passing score (see Geisinger & McCormick, 2010). A state may accept the recommended

[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 recommened passing scores for each panel are presented.

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 Agriculture 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 Agriculture 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 AGRICULTURE TEST

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

The two-hour assessment contains 120 selected response items[4] covering seven content areas: *Agribusiness Systems* (approximately 14 items), *Animal Systems* (approximately 20 items), *Food Science and Biotechnology Systems* (approximately 14 items), *Environmental and Natural Resource Systems* (approximately 16 items), *Plant Systems* (approximately 20 items), *Power, Structural, and Technical Systems* (approximately 18 items), and *Leadership and Career Development* (approximately 18 items).[5] The reporting scale for the Praxis Agriculture 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 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] Ten of the 120 multiple-choice 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 Agriculture 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 28 educators representing 18 states. (See Appendix A for a listing of panelists.) Eighteen panelists were teachers, and ten were college faculty. Nine of the ten faculty members' job responsibilities included the training of agriculture 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	18	64%
College faculty	10	36%
Race		
White	28	100%
Gender		
Female	10	36%
Male	18	64%
Are you currently certified to teach this subject in your state?		
Yes	24	86%
No	4	14%
Are you currently teaching this subject in your state?		
Yes	23	82%
No	5	18%
Are you currently supervising or mentoring other teachers of this subject?		
Yes	15	54%
No	13	46%
At what K–12 grade level are you currently teaching this subject?		
Middle school (6–8 or 7–9)	3	11%
High school (9–12 or 10–12)	8	29%
Middle and High school	7	25%
Not currently teaching at the K–12 level	10	36%

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

	N	%
Including this year, how many years of experience do you have teaching this subject?		
3 years or less	4	14%
4–7 years	10	36%
8–11 years	7	25%
12–15 years	3	11%
16 years or more	4	14%
Which best describes the location of your K–12 school?		
Urban	1	4%
Suburban	4	14%
Rural	13	46%
Not currently working at the K–12 level	10	36%
If you are college faculty, are you currently involved in the training/preparation of teacher candidates in this subject?		
Yes	9	32%
No	1	4%
Not college faculty	18	64%

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	62.86	63.20
Lowest	57.50	53.75
Highest	75.80	74.60
SD	4.84	6.12
SEJ	1.25	1.70

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 Agriculture test are 62.86 for Panel 1 and 63.20 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 — 63 for Panel 1 and 64 Panel 2. The scaled scores associated with 63 and 64 raw points are 146 and 147, 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 Agriculture test is 63.03 (out of a possible 110 raw-score points). The value was rounded to 64 (next highest raw score) to determine the functional recommended passing score. The scaled score associated with 64 raw points is 147.

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

Table 4

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

Recommended passing score (CSEM)		Scale score equivalent
64 (5.20)		147
-2 CSEMs	54	134
-1 CSEM	59	141
+ 1 CSEM	70	155
+ 2 CSEMs	75	162

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

FINAL EVALUATIONS

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

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

All panelists *strongly agreed* or *agreed* that they understood the purpose of the study and that 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 of the panelists indicated they were at least *somewhat comfortable* with the passing score they recommended; 17 of the 28 panelists were *very comfortable*. Twenty-seven of the 28 panelists indicated the recommended passing score was *about right* with the remaining panelist indicating 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 Agriculture 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 Agriculture test, the recommended passing score[8] is 64 out of a possible 110 raw-score points. The scaled score associated with a raw score of 64 is 147 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: Agriculture (5701)*. 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**
Brent Arndt	North Valley Career and Tech Center/Park River (ND)
Harry Boone	West Virginia University (WV)
Mark Breeding	W. T. Chipman (DE)
Larry Butler	Grant County High School (KY)
Jeremy Carkuff	Tioga High School (ND)
Josh Dahlem	Stanley High School (LA)
Don Edgar	University of Arkansas (AR)
James Graham	University of WI – River Falls (WI)
Randall J. Haefele	HEM High School – Carbon County SD #2 (WY)
Emily Hester	Westminster High School (MD)
Kathleen M. Jones	Juniata College (PA)
Shannon G. Lawrence	Clemson University (SC)
Rebecca G. Lawver	Utah State University (UT)
Jon Lechtenberg	Southern Jr/Sr High School (NE)
Amanda Levzow-Seichter	Wisconsin Dells High School (WI)
Tracy Marchini	Eastern Lancaster County School District (PA)
Bart Mattingly	Washington County Schools (KY)
Tonya Mortensen	Medicine Valley Public School (NE)
Brynn Mulvihill	Everett Meredith Middle School (DE)
Tiffany Myers	Spanish Fork Junior High School (UT)
Thomas H. Paulsen	Iowa State University (IA)
Caleb Plyler	Blevins High School (AR)
Karen Roudabush	Bridgewater-Emery School District (SD)
Scott Smalley	South Dakota State University (SD)
Christopher Stripling	University of Tennessee (TN)
Caroline Tucker	North Greene High School (TN)
Shannon Washburn	Kansas State University (KS)
Wesley Wilson	Pahranagat Valley High/Middle School (NV)

APPENDIX B

STUDY AGENDA

AGENDA

Praxis Agriculture (5701)
Standard-Setting Study

Day 1

Welcome and Introduction

Overview of Standard Setting and the Praxis Agriculture Test

Review the Praxis Agriculture Test

Discuss the Praxis Agriculture Test

Lunch

Define the Knowledge/Skills of a Target Candidate

Break

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

Lunch

Feedback on Round 2 Recommended Cut Score

Complete Final Evaluation

Collect Materials; End of Study

APPENDIX C

TARGET CANDIDATE DESCRIPTION

Description of the Target Candidate[9]

A target candidate …

I. AGRIBUSINESS SYSTEMS

1. Is familiar with the fundamentals of saving, investment, credit and marketing principles needed to accomplish agribusiness objectives
2. Knows the record keeping practices needed to accomplish agribusiness objectives (asset, liabilities, income expense, net worth, balance sheet)

II. ANIMAL SYSTEMS

3. Knows basic animal systems (reproductive, skeletal, digestive)
4. Is familiar with basic production (beef, poultry and dairy) and basic management structures and principles (nutrition, care, safety)

III. FOOD SCIENCE AND BIOTECHNOLOGY SYSTEMS

5. Is familiar with the history, major issues, current trends and safe laboratory procedures and the application of biotechnology in the food products and processing industry

IV. ENVIRONMENTAL AND NATURAL RESOURCE SYSTEMS

6. Is familiar with the land use practices, major natural cycles and their effect on the environment and natural resources
7. Is familiar with use, production, and processing of natural resources and their impact on conventional and alternate energy sources

V. PLANT SYSTEMS

8. Knows the basic principles of general safety issues as related to plant systems and production
9. Knows the basic principles of identification (woody, leaf shapes, root system), classification (monocot, dicot), anatomy (stem, internode) and physiology (photosynthesis) as related to plant production and management
10. Is familiar with the basic horticultural and agronomic production and management practices such as:
    a. Propagation, cultivation, harvesting
    b. Characteristics of soils/growing media

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

Description of the Target Candidate[10] (continued)

A target candidate …

VI. POWER, STRUCTURAL, AND TECHNICAL SYSTEMS

11. Is familiar with the basic principles of electricity, power sources, small engines, metal fabrication, construction and welding including applied math
12. Understands the safe operation of commonly found shop equipment
13. Knows maintenance of equipment and proper laboratory management, including the storage of potentially hazardous materials
14. Is familiar with the application of technology to agriculture industry

VII. LEADERSHIP AND CAREER DEVELOPMENT

15. Understands the relationship between local program planning, supervised agricultural experiences and the national FFA organization
16. Knows the principles of leadership and effective communication skills, such as:
    a. Individual and team leadership
    b. Research skills to make informed decisions
17. Knows the foundational skills (work ethics, resume writing, interview) of career development across the various pathways of agriculture

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

APPENDIX D

RESULTS

Table D1

Panel Member Demographics (by Panel)

	Panel 1		Panel 2	
	N	**%**	***N***	**%**
Current position				
Teacher	11	73%	7	54%
College faculty	4	27%	6	46%
Race				
White	15	100%	13	100%
Gender				
Female	7	47%	3	23%
Male	8	53%	10	77%
Are you currently certified to teach this subject in your state?				
Yes	13	87%	11	85%
No	2	13%	2	15%
Are you currently teaching this subject in your state?				
Yes	13	87%	10	77%
No	2	13%	3	23%
Are you currently supervising or mentoring other teachers of this subject?				
Yes	9	60%	6	46%
No	6	40%	7	54%
At what K–12 grade level are you currently teaching this subject?				
Middle school (6–8 or 7–9)	1	7%	2	15%
High school (9–12 or 10–12)	5	33%	3	23%
Middle and High School	5	33%	2	15%
Not currently teaching at the K–12 level	4	27%	6	46%

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	2	13%	2	15%
4–7 years	6	40%	4	31%
8–11 years	6	40%	1	8%
12–15 years	0	0%	3	23%
16 years or more	1	7%	3	23%
Which best describes the location of your K–12 school?				
Urban	0	0%	1	8%
Suburban	3	20%	1	8%
Rural	8	53%	5	38%
Not currently working at the K–12 level	4	27%	6	46%
If you are college faculty, are you currently involved in the training/preparation of teacher candidates in this subject?				
Yes	4	27%	5	38%
No	0	0%	1	8%
Not college faculty	11	73%	7	54%

Table D2

Passing Score Summary by Round of Judgments

	Panel 1		Panel 2	
Panelist	**Round 1**	**Round 2**	**Round 1**	**Round 2**
1	57.95	57.50	71.05	63.10
2	60.65	60.95	60.30	62.30
3	60.40	60.50	63.50	62.40
4	67.80	65.90	60.90	62.70
5	64.75	65.65	58.10	60.80
6	61.95	61.30	76.40	74.60
7	63.60	64.00	46.75	56.30
8	63.80	64.00	55.65	58.85
9	79.50	75.80	72.85	69.90
10	59.75	62.70	71.65	71.90
11	58.80	58.60	56.60	58.40
12	68.75	68.75	51.20	53.75
13	55.10	57.50	69.35	66.65
14	59.00	60.70		
15	57.90	59.00		
Average	62.65	62.86	62.64	63.20
Lowest	55.10	57.50	46.75	53.75
Highest	79.50	75.80	76.40	74.60
SD	5.99	4.84	9.08	6.12
SEJ	1.55	1.25	2.52	1.70

Table D3

Final Evaluation: Panel 1

	Strongly agree		Agree		Disagree		Strongly disagree	
	N	**%**	***N***	**%**	***N***	**%**	***N***	**%**
• I understood the purpose of this study.	14	93%	1	7%	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.	14	93%	1	7%	0	0%	0	0%
• The process of making the standard-setting judgments was easy to follow.	15	100%	0	0%	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 target candidate	14	93%	1	7%	0	0%
• The between-round discussions	7	47%	8	53%	0	0%
• The knowledge/skills required to answer each test item	9	60%	6	40%	0	0%
• The passing scores of other panel members	0	0%	14	93%	1	7%
• My own professional experience	6	40%	8	53%	1	7%

	Very comfortable		Somewhat comfortable		Somewhat uncomfortable		Very uncomfortable	
	N	%	*N*	%	*N*	%	*N*	%
• Overall, how comfortable are you with the panel's recommended passing score?	11	73%	4	27%	0	0%	0	0%

	Too low		About right		Too high	
	N	%	*N*	%	*N*	%
• Overall, the recommended passing score is:	1	7%	14	93%	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.	12	92%	1	8%	0	0%	0	0%
• The instructions and explanations provided by the facilitator were clear.	9	69%	4	31%	0	0%	0	0%
• The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	7	54%	6	46%	0	0%	0	0%
• The explanation of how the recommended passing score is computed was clear.	8	62%	5	38%	0	0%	0	0%
• The opportunity for feedback and discussion between rounds was helpful.	12	92%	1	8%	0	0%	0	0%
• The process of making the standard-setting judgments was easy to follow.	7	54%	6	46%	0	0%	0	0%

Table D4 (continued)

Final Evaluation: Panel 2

How influential was each of the following factors in guiding your standard-setting judgments?	Very influential		Somewhat influential		Not influential	
	N	%	*N*	%	*N*	%
• The description of the target candidate	12	92%	1	8%	0	0%
• The between-round discussions	8	62%	5	38%	0	0%
• The knowledge/skills required to answer each test item	9	69%	4	31%	0	0%
• The passing scores of other panel members	0	0%	12	92%	1	8%
• My own professional experience	4	31%	8	62%	1	8%

	Very comfortable		Somewhat comfortable		Somewhat uncomfortable		Very uncomfortable	
	N	%	*N*	%	*N*	%	*N*	%
• Overall, how comfortable are you with the panel's recommended passing score?	6	46%	7	54%	0	0%	0	0%

	Too low		About right		Too high	
	N	%	*N*	%	*N*	%
• Overall, the recommended passing score is:	0	0%	13	100%	0	0%