



Indiana Department of Education

Dr. Katie Jenner, Secretary of Education

To: State Board of Education

From: John Keller, Chief Technology Officer
Risa A. Regnier, Director Educator Licensing

Date: March 3, 2021

Re: Final approval of five new Praxis licensure tests with cut scores

Background

In October 2018, the IDOE released an RFP for a new licensure test vendor and program as required by the 2018 General Assembly. In July 2019, in accordance with the results of the RFP and action by the 2019 General Assembly, Educational Testing Service (ETS)--with its nationally available Praxis test program--was adopted by the SBOE to provide educator licensure tests as of September 1, 2021.

Since its adoption in July 2019, ETS and IDOE began working to establish test cut scores for the existing Praxis tests Indiana will use, and the development of several new Praxis tests that will be used in Indiana and made available to other ETS client states. As of December 2020, 42 existing Praxis tests and cut scores had been approved by the board. That level of test and cut score adoption has allowed the DOE to work with the two test vendors to facilitate an earlier transition to many of the Praxis tests than previously planned. The 42 tests and scores already adopted will be available for test administration after July 1, 2021, rather than September 1, 2021.

Included in the newly developed Praxis tests are the two Early Childhood Generalist subtests in STEM and Humanities, the two Elementary Generalist subtests in STEM and Humanities, and the P-12 pedagogy test. Following development of these tests, a multi-state standard setting (cut score) study was conducted that included Indiana teachers as participants. The tests and recommended cut scores were posted for public comment for 30 days. The comments are summarized below:

- There were no public comments for the Early Childhood Generalist subtests.
- Three public comments were received for the Elementary Generalist subtests indicating disagreement with the recommended cut scores: 1) no licensure tests should be required; 2) the cut scores should be “on the curve;” and 3) a comment asking about the weighting of the test items.
- One comment was received for the P-12 pedagogy test indicating disagreement with the recommended score, stating that the score should be “on the curve” and ongoing PD should be required for teachers.



Indiana Department of Education

Dr. Katie Jenner, Secretary of Education

The technical reports describing the new Praxis tests, the multi-state standard setting process, and the recommended cut scores for Indiana accompany this memo.

Action Needed

Under 511 IAC 15-7-2 the board must vote final approval of licensure tests and cut scores after 30 days of public comment.

DOE Recommendation

Final approval of the newly developed Praxis STEM and Humanities subtests for Early Childhood Generalist and Elementary Generalist, the Praxis test for P-12 pedagogy, and the recommended cut scores for each of those five tests.



Multistate Standard-Setting Technical Report

**PRAXIS[®] EARLY CHILDHOOD ASSESSMENT (5026):
READING AND LANGUAGE ARTS AND SOCIAL STUDIES SUBTEST
(5027)
MATHEMATICS AND SCIENCE SUBTEST (5028)**

Licensure and Credentialing Research

ETS

Princeton, New Jersey

January 2021

EXECUTIVE SUMMARY

To support the decision-making process of education agencies establishing a passing score (cut score) for the *Praxis*[®] Early Childhood Assessment (5026), research staff from Educational Testing Service (ETS) designed and conducted a distance-based multistate standard-setting study.

PARTICIPATING STATES

Panelists from nine states and Washington, D.C., were recommended by their respective education agencies. The education agencies recommended panelists with (a) experience as either early childhood teachers or college faculty who prepare early childhood teachers and (b) familiarity with the knowledge and skills required of beginning early childhood teachers.

RECOMMENDED PASSING SCORE

ETS provides a multistate standard-setting study technical report, with the recommended passing score, to help education agencies determine appropriate operational passing scores for the two subtests of the Praxis Early Childhood Assessment. The panelists completed the standard-setting process and recommended two scores, one for each subtest—Reading and Language Arts and Social Studies (5027) and Mathematics and Science (5028). The recommended passing scores are listed in Table 1. The scale scores associated with each of the raw score recommendations are on a 100–200 scale.

Table 1.
Recommended Passing Scores (raw values and scale scores)

Subtest Title	Test Code	Recommended raw score	Total raw scores	Praxis scale score
Reading and Language Arts and Social Studies	5027	51	75	161
Mathematics and Science	5028	50	75	160

INTRODUCTION

To support the decision-making process for education agencies establishing passing scores (cut scores) for the *Praxis*[®] Early Childhood Assessment (5026), research staff from Educational Testing Service (ETS) designed and conducted a distance-based multistate standard-setting study in December 2020. Education agencies¹ recommended panelists with (a) experience as either early childhood teachers or college faculty who prepare early childhood teachers and (b) familiarity with the knowledge and skills required of beginning early childhood teachers. Nine states and Washington, D.C. (Table 2) were represented by 25 panelists. (See Appendix A for the names and affiliations of the panelists.)

Table 2.
Participating States, Washington, D.C., and Number of Panelists

Alaska (2 panelists)	Nevada (3 panelists)
Arkansas (2 panelists)	New Mexico (5 panelists)
Delaware (1 panelist)	Washington, D.C. (3 panelists)
Hawaii (1 panelist)	West Virginia (4 panelists)
Indiana (4 panelists)	

The following technical report contains three sections. The first section describes the content and format of the test. The second section describes the standard-setting processes and methods. The third section presents the results of the standard-setting study.

ETS provides a recommended passing score from the multistate standard-setting study to education agencies. In each state and D.C., the department of education, the board of education, or a designated educator licensure board is responsible for establishing the operational passing score in accordance with applicable regulations. This study provides a recommended passing score,² which represents the combined judgments of two panels of experienced educators. Each state and D.C., may want to consider the recommended passing score but also other sources of information when setting the final *Praxis* Early Childhood Assessment passing score (see Geisinger & McCormick, 2010). A state and D.C., may accept the recommended passing score, adjust the score upward to reflect more stringent expectations, or adjust the score downward to reflect more lenient expectations. There is no *correct*

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

² In addition to the recommended passing score averaged across the two panels, the recommended passing scores for each panel are presented.

decision; the appropriateness of any adjustment may only be evaluated in terms of its meeting the state and D.C.'s needs.

Two sources of information to consider when setting the passing score are the standard error of measurement (SEM) and the standard error of judgment (SEJ). The former addresses the reliability of the *Praxis* Early Childhood Assessment score and the latter, the reliability of panelists' passing-score recommendation. The SEM allows a state and D.C. to recognize that any test score on any standardized test—including a *Praxis* Early Childhood Assessment score—is not perfectly reliable. A test score only *approximates* what a candidate truly knows or truly can do on the test. The SEM, therefore, addresses the question: How close of an approximation is the test score to the *true* score? The SEJ allows a state and D.C. to gauge the likelihood that the recommended passing score from a particular panel would be similar to the passing scores recommended by other panels of experts similar in composition and experience. The smaller the SEJ, the more likely that another panel would recommend a passing score consistent with the recommended passing score. The larger the SEJ, the less likely the recommended passing score would be reproduced by another panel.

In addition to measurement error metrics (e.g., SEM, SEJ), each state and D.C. should consider the likelihood of classification errors. That is, when adjusting a passing score, policymakers should consider whether it is more important to minimize a false-positive decision or to minimize a false-negative decision. A false-positive decision occurs when a candidate's test score suggests that he should receive a license/certificate, but his actual level of knowledge/skills indicates otherwise (i.e., the candidate does not possess the required knowledge/skills). A false-negative decision occurs when a candidate's test score suggests that she should not receive a license/certificate, but she actually does possess the required knowledge/skills. The state and D.C. needs to consider which decision error is more important to minimize.

OVERVIEW OF THE *PRAXIS*[®] EARLY CHILDHOOD ASSESSMENT

The *Praxis*[®] Early Childhood *Study Companion* document (ETS, in press) describes the purpose and structure of the test. In brief, the test measures whether entry-level early childhood teachers have the knowledge/skills believed necessary for competent professional practice.

Candidates must pass both subtests in order to pass the *Praxis* Early Childhood Assessment. They may retake the subtest that they did not pass in order to fulfill the requirement. The reporting scale for each subtest ranges from 100 to 200.

READING AND LANGUAGE ARTS AND SOCIAL STUDIES (5027)

This 2-hour subtest contains items 90 selected-response items covering two content areas: *Reading and Language Arts* (approximately 60 items) and *Social Studies* (approximately 30 items).³ The reporting scale for the *Praxis* Reading and Language Arts and Social Studies subtest ranges from 100 to 200 scale-score points.

MATHEMATICS AND SCIENCE (5028)

This 2-hour subtest contains items 90 selected-response items covering two content areas: *Mathematics* (approximately 58 items) and *Science* (approximately 32 items).³ The reporting scale for the *Praxis* Mathematics and Science subtest ranges from 100 to 200 scale-score points.

PROCESSES AND METHODS

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

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

³ The number of items for each content area may vary slightly from form to form of the test.

REVIEWING THE TEST

Test familiarization was the first activity for the panel. The purpose of test familiarization is for the panelists to review the test and become familiar with the manner in which a candidate would take the test. After the facilitator described the purpose of the review and how to access the test⁴, the standard-setting panelists took the test and had a discussion of the content measured. This discussion helped bring the panelists to a shared understanding of what the test measures.

The test discussion covered the major content areas being addressed by the test. Panelists were asked to remark on any content areas that would be particularly challenging for entry-level teachers or areas that address content particularly important for entry-level teachers. Overall, this discussion serves to reduce potential judgment errors later in the standard-setting process.

For this two-panel, multistate standard-setting study, the first test that was reviewed differed for each panel. The educators on Panel 1 first viewed the Reading and Language Arts (RLA) and Social Studies subtest. The educators on Panel 2 reviewed the Mathematics (Math) and Science subtest as their first activity. The facilitator took notes of each panel's content discussion. After completing the entire standard-setting process on the first subtest, each panel would then switch to the other subtest. The notes from the content discussion are provided to help each panel see the comparisons in the discussion points.

DESCRIBING THE JUST QUALIFIED CANDIDATE

Following the review of the test, panelists worked together to describe the just qualified candidate—specific to the content measured on the subtest the panel reviewed. The *just qualified candidate description* plays a central role in standard setting (Perie, 2008); the goal of the standard-setting process is to identify the test score that aligns with this description.

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

The final 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*

⁴ The computer-administered test items were available through the ETS IBIS Content Review Tool.

qualified candidate. The typed description was made available for panelists to download and use during subsequent phases of the study (see Appendix C for the just qualified candidate descriptions).

Panel 1 created the just qualified candidate description, specific to the knowledge and skills measured on the RLA and Social Studies subtest. Panel 2 created the just qualified candidate description for the Math and Science subtest. Each panel completed their test familiarization for the second subtest, aided by the discussion notes from the other panel. Afterwards, they worked to understand the just qualified candidate description of that second subtest. So, Panel 1 received the description created by Panel 2 for the Math and Science subtest; Panel 2 received the description created by Panel 1 for the RLA and Social Studies subtest.

To *understand* the just qualified candidate description provided by the other panel, each set of panelists worked together to create performance indicators—observable, or measurable, actions that a candidate could do to demonstrate the level of knowledge and skills described in the description. The panel would discuss these statements, specifying what the just qualified candidate *could do* and what the candidate *could not do*.

PANELISTS' JUDGMENTS

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

Panelists were asked to approach the judgment process in two stages. First, they reviewed both the description of the just qualified candidate and the item and determined what was the probability that the just qualified candidate would answer the question correctly. The facilitator encouraged the panelists to consider the following rules of thumb to guide their decision:

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

- Items in the .70 to 1 range were those that the just qualified candidate would have a high chance of answering correctly.

Next, panelists decided how to refine their judgment within the range. For example, if a panelist thought that there was a high chance that the just qualified candidate would answer the question correctly, the initial decision would be in the .70 to 1 range. The second decision for the panelist was to judge if the likelihood of answering it correctly is .70, .80, .90, .95 or 1.

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

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

The panelists discussed their item-level judgments. These discussions helped panelists maintain a shared understanding of the knowledge/skills of the just qualified candidate and helped to clarify aspects of items that might not have been clear to all panelists during the Round 1 judgments. The purpose of the discussion was not to encourage panelists to conform to another's judgment, but to understand the different relevant perspectives among the panelists.

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

Each panel completed both rounds of judgments on the first subtest before starting over with the test familiarization for the second subtest, continuing the process and completing both rounds of judgments. Other than content discussion notes and 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 3 presents a summary of the panelists' demographic information. The panel included 25 educators representing nine states and D.C.. (See Appendix A for a listing of panelists.) Seven panelists were teachers, 14 were college faculty, one was an administrator or department head, and three held other positions. All 14 faculty members' job responsibilities included the training of early childhood teachers. The number of experts by panel and their demographic information are presented in Appendix D (Table D1).

Table 3.
Panel Member Demographics (Across Panels)

	<i>N</i>	<i>%</i>
Current position		
Teacher	7	28
Administrator/Department head	1	4
College faculty	14	56
Other	3	12
Race		
White	20	80
Black or African American	2	8
Hispanic or Latino	1	4
Asian or Asian American	2	8
Gender		
Female	25	100
Male	0	0
Are you currently certified to teach early childhood?		
Yes	12	48
No	13	52
Are you currently teaching early childhood in your state?		
Yes	13	52
No	12	48
Are you currently supervising or mentoring other teachers of early childhood?		
Yes	20	80
No	5	20

(table continues)

Table 3. (continued)
Panel Member Demographics (Across Panels)

	<i>N</i>	<i>%</i>
At what grade level(s) are you currently teaching early childhood? (Select all that apply)		
Birth to pre-kindergarten	3	12
Pre-kindergarten	9	36
Kindergarten	7	28
Grade 1	4	16
Grade 2	3	12
Grade 3	2	8
Other	16	64
Including this year, how many years of experience do you have teaching early childhood?		
3 years or less	2	8
4–7 years	4	16
8–11 years	5	20
12–15 years	5	20
16 years or more	9	36
Which best describes the location of your P–12 school?		
Urban	6	24
Suburban	1	4
Rural	3	12
Not currently working at the P–12 level	15	60
Considering the current and the past three school years, which content area(s) do you primarily teach? (Select all that apply)		
I do not teach in the P-12 grade levels.	15	60
Mathematics	10	40
English Language Arts & Reading	10	40
Social Studies	9	36
Science	9	36
Other	11	44
If you are college faculty, are you currently involved in the training/preparation of teacher candidates in early childhood?		
Yes	14	56
No	0	0
Not college faculty	11	44

STANDARD-SETTING JUDGMENTS

Tables 4 and 5 summarize 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.⁵ It indicates how likely it would be for several other panels of educators similar in makeup, experience, and standard-setting training to the current panel to recommend the same passing score on the same form of the test. 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 (Tables D2 and D3).

Table 4.

Reading and Language Arts and Social Studies Summary of Round 2 Standard-setting Judgments Per Panel

Test Code		Panel 1	Panel 2
5027	Average	52.36	49.25
	Lowest	44.50	46.50
	Highest	64.00	53.90
	SD	6.03	2.08
	SEJ	1.74	0.58

Table 5.

Mathematics and Science Summary of Round 2 Standard-setting Judgments Per Panel

Test Code		Panel 1	Panel 2
5028	Average	53.01	46.93
	Lowest	43.60	40.40
	Highest	64.80	53.90
	SD	6.68	3.62
	SEJ	1.93	1.00

Round 1 judgments are made without discussion among the panelists. The most variability in judgments, therefore, is typically present in the first round. Round 2 judgments, however, are informed by panel discussion; thus, it is common to see a decrease both in the standard deviation and SEJ. This

⁵ An SEJ assumes that panelists are randomly selected and that standard-setting judgments are independent. It is seldom the case that panelists are randomly sampled, and only the first round of judgments may be considered independent. The SEJ, therefore, likely underestimates the uncertainty of passing scores (Tannenbaum & Katz, 2013).

decrease — indicating convergence among the panelists’ judgments — was observed for each panel (see Tables D2 and D3 in Appendix D). The Round 2 average score is the panel’s recommended passing score.

The panels’ passing score recommendations for the *Praxis* Early Childhood Assessment are listed in Table 6. The values were rounded to the next highest whole number, to determine the functional recommended passing scores. There are 75 raw score points available for the Reading and Language Arts and Social Studies subtest (5027). For the Mathematics and Science subtest (5028) there are 75 available raw score points.

Table 6.
Recommended Passing Scores per Panel and Subtest

Subtest: Panel	Recommended Score		Praxis Score (100 – 200 Scale)
	Raw	Rounded	
RLA & Social Studies 5027: Panel 1	52.36	53	165
RLA & Social Studies 5027: Panel 2	49.25	50	160
Math & Science 5028: Panel 1	53.01	54	168
Math & Science 5028: Panel 2	46.93	47	155

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. Table 7 displays those results.

Table 7.
Recommended Passing Scores per Subtest (Panels Scores Combined)

Subtest	Panels’ Averaged Recommendations		Praxis Score (100 – 200 Scale)
	Raw	Rounded	
RLA and Social Studies 5027	50.81	51	161
Math and Science 5028	49.97	50	160

Tables 8 and 9 present the estimated conditional standard error of measurement (CSEM) around the recommended passing score. A standard error represents the uncertainty associated with a test score. The scale scores associated with one and two CSEM above and below the recommended passing score are provided. The conditional standard error of measurement provided is an estimate.

Table 8.
Reading and Language Arts and Social Studies Passing Scores Within 1 and 2 CSEM of the Recommended Passing Score⁶

	Recommended passing score (CSEM)	Scale score equivalent
	51 (4.07)	161
Test: 5027	-2 CSEM	43
	-1 CSEM	47
	+ 1 CSEM	56
	+ 2 CSEM	60
		179

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

Table 9.
Mathematics and Science Passing Scores Within 1 and 2 CSEM of the Recommended Passing Score⁶

	Recommended passing score (CSEM)	Scale score equivalent
	50 (4.11)	160
Test: 5028	-2 CSEM	42
	-1 CSEM	46
	+ 1 CSEM	55
	+ 2 CSEM	59
		177

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

FINAL EVALUATIONS

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

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

All panelists *strongly agreed* or *agreed* that they understood the purpose of the study. Twenty-three of the 25 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-

⁶ The unrounded CSEM value is added to or subtracted from the rounded passing-score recommendation. The resulting values are rounded up to the next-highest whole number and the rounded values are converted to scale scores.

setting judgments. All panelists *strongly agreed* or *agreed* that the standard-setting process was easy to follow.

All but one of the panelists reported that the description of the just qualified candidate was at least *somewhat influential* in guiding their standard-setting judgments; 19 of the 25 panelists indicated the description was *very influential*. All but one of the panelists reported that between-round discussions were at least *somewhat influential* in guiding their judgments. More than two-thirds of the panelists (22 of the 25 panelists) indicated that their own professional experience was *very influential* in guiding their judgments.

For the reading and language arts and social studies subtest, all of the panelists indicated they were at least *somewhat comfortable* with the passing score they recommended; 18 of the 25 panelists were *very comfortable*. Twenty-four of the 25 panelists indicated the recommended passing score was *about right* with the remaining panelist indicating that the passing score was *too low*.

For the mathematics and science subtest, all of the panelists indicated they were at least *somewhat comfortable* with the passing score they recommended; 19 of the 25 panelists were *very comfortable*. Twenty-four of the 25 panelists indicated the recommended passing score was *about right* with the remaining panelist indicating that the passing score was *too high*.

SUMMARY

To support the decision-making process for education agencies establishing a passing score (cut score) for the *Praxis* Early Childhood Assessment, 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 appropriate operational passing scores. The panelists from each of the two panels completed the standard-setting process and recommended⁷ two scores—one for each of the two subjects. The recommended passing scores are listed in Table 10. The scale scores listed associated with each of the raw score recommendations are on a 100–200 scale.

Table 10.
Recommended Passing Scores (raw values and scale scores)

Subtest Title	Test Code	Recommended raw score	Total raw scores	Praxis scale score
Reading and Language Arts and Social Studies	5027	51	75	161
Mathematics and Science	5028	50	75	160

⁷ Results from the two panels participating in the study were averaged to produce the recommended passing score.

REFERENCES

- Brandon, P. R. (2004). Conclusions about frequently studied modified Angoff standard-setting topics. *Applied Measurement in Education, 17*, 59–88.
- ETS. (in press). *The Praxis Series®: The Praxis Study Companion: Early Childhood (5026)*. 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>
Brittany Behenna Griffith	Santa Fe Public Schools (NM)
Tammy Benson	University of Central Arkansas (AR)
Renee Boyer	Indiana Wesleyan University (IN)
Youngae Choi	Nevada State College (NV)
Robyn Chun	University of Hawaii at Manoa (HI)
Elise Colomb	University of Delaware (DE)
Mary Ellen Dannenberg	Turquoise Trail Charter School (NM)
Cathy Gonzales	Oakland City University (IN)
Laura Herold	University of Arkansas (AR)
Teri Hogg	Purdue University Fort Wayne (IN)
Dancey Howes	Webster County Schools (WV)
Amia Johnson	Early Childhood Academy PCS (DC)
Jennifer Lee	Trinity University (DC)
Paula Lucas	Marshall University (WV)
Sarah Massie	District of Columbia Public Schools (DC)
Marissa Persinger	Eastbrook Elementary School (WV)
Linda Pheasant	Clark County School District (NV)
Veronica Plumb	University of Alaska Fairbanks (AK)
Hilary Seitz	University of Alaska Anchorage (AK)
Jackie Shipley	Central New Mexico Community College (NM)
Stephanie Stopiak	West Virginia University at Parkersburg (WV)
JoAnna Valencia	Bernalillo Public Schools (NM)

(table continues)

Participating Panelists With Affiliation (continued)

<u>Panelist</u>	<u>Affiliation</u>
Sarah Vogt	Central New Mexico Community College (NM)
Cayla Wise	University of Saint Francis (IN)
Kristin Withey	The Meadows School (NV)

APPENDIX B
STUDY AGENDA

AGENDA

***Praxis*[®] Early Childhood Assessment (5026) Standard-Setting Study**

Day 1

Welcome and Introduction

Standard-setting overview presentation

- Q&A about the training

Test familiarization for the first subtest

- Panel 1: Reading & Language Arts (RLA) & Social Studies (5027) subtest
- Panel 2: Mathematics & Science (5028) subtest

Break

Discuss the content measured on the first subtest

Lunch Break

Overview of the Just Qualified Candidate

Define the Just Qualified Candidate for the subtest

- Panel 1: Reading & Language Arts (RLA) & Social Studies (5027) subtest
- Panel 2: Mathematicaic and Science (5028) subtest

Break

Define the Just Qualified Candidate for the subtest (continued)

Standard-setting training presentation

Practice Round: Selected-response standard-setting judgments

End of Day 1

AGENDA

***Praxis*[®] Early Childhood Assessment (5026) Standard-Setting Study**

Day 2

Overview of the Day

Practice Round: Data Discussion

Evaluation of the standard-setting training (poll)

Round 1: Selected-response standard-setting judgments

Break

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

Lunch Break

Round 1 feedback: Summary data

Evaluation of data presentation (poll)

Round 1 feedback: Item-level data and Round 2 judgments

Break

Round 1 feedback: Item-level data and Round 2 judgments
(*continued*)

Test familiarization for the second subtest

- Panel 1: Mathematics & Science (5028)

Panel 2: Reading Language Arts (RLA) & Social Studies (5027)

Wrap up and Complete Day 2

AGENDA

***Praxis*[®] Early Childhood Assessment (5026) Standard-Setting Study**

Day 3

Overview of the Day

Discuss the content measured on the second subtest

Reviewing the Just-Qualified Candidate – Whole group discussion

- Panel 1: Mathematics & Science (5028) subtest
- Panel 2: Reading & Language Arts (RLA) & Social Studies (5027) subtest

Break

Round 1: Selected-response standard-setting judgments

Lunch Break

Round 1 Feedback Summary data

Evaluation of the data presentation (poll)

Round 1 Feedback and Round 2 Judgments

Break

Feedback on Round 2 Recommended Cut Scores

Complete Final Evaluation

End of Study

APPENDIX C

JUST QUALIFIED CANDIDATE DESCRIPTION

Description of the Just Qualified Candidate⁸

A just qualified candidate...

Reading and Language Arts

1. Understands the complexity of the five components (Phonics, phonemic awareness, fluency, comprehension, vocabulary) of reading literacy
2. Understands oral language (e.g. articulation, awareness of audience, eye contact, volume, collaborative discussions, and phonological awareness)
3. Understands the components of writing (e.g., stages, types, cycle, digital media, and research)
4. Is familiar with the integration of literacy across content areas and connects to the diverse lived experiences of all children (e.g., literature, background)
5. Is familiar with English Language acquisition and the impact on English literacy development.
6. Understanding the components of literature and informational texts such as authors purpose, inferencing, themes, cause & effect, and structural elements.

Social Studies

7. Knows how people of different cultural backgrounds interact with their environment, self, family, neighborhood and organizations
8. Knows general content of geography and understands its impact on physical and human systems
9. Knows essential concepts of personal identity, cultural competence (responsiveness), civic participation in the environment and being able to integrate these across the curriculum and within the classroom environment.
10. Understands chronological thinking skills and historical data (e.g., time lines, maps, graphs, and tables)

⁸ Description of the just qualified candidate focuses on the knowledge/skills that differentiate a *just* from a *not quite* qualified candidate.

Description of the Just Qualified Candidate (continued)⁹

A just qualified candidate...

Mathematics

1. Recognizes appropriate levels of student skill development—and understands concepts at a basic level—in the mathematical domains of one to one correspondence, number sense, operations in base ten, concepts related to fractions, algebraic thinking, geometry, measurement, data analysis and categorizing
2. Recognizes the different ways of showing and analyzing mathematical knowledge through concrete representations (such as manipulatives) as well as abstract representations
3. Recognizes opportunities to make connections by applying basic mathematical knowledge to real life experiences.

Science

4. Recognizes the basic positive and negative influences of science and technology (past and modern) on the environment and society.
5. Has basic knowledge of scientific content and practices (e.g., inquiry methods), across all scientific subdisciplines: physical science, earth and space science, life science, and engineering/technology.
6. Recognizes opportunities to make connections by applying basic scientific knowledge from the science subdisciplines to real life experiences.

Both Mathematics and Science

7. Recognizes opportunities for (i) integration of literacy across the content areas and (ii) making connections to the experiences and backgrounds of all children

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	
	<i>N</i>	%	<i>N</i>	%
Current position				
Teacher	3	25	4	31
Administrator/Department head	0	0	1	8
College faculty	7	58	7	54
Other	2	17	1	8
Race				
White	10	83	10	77
Black or African American	1	8	1	8
Hispanic or Latino	0	0	1	8
Asian or Asian American	1	8	1	8
Gender				
Female	12	100	13	100
Male	0	0	0	0
Are you currently certified to teach early childhood?				
Yes	6	50	6	46
No	6	50	7	54
Are you currently teaching early childhood in your state?				
Yes	6	50	7	54
No	6	50	6	46
Are you currently supervising or mentoring other teachers of early childhood?				
Yes	11	92	9	69
No	1	8	4	31
At what P-12 grade levels are you currently teaching early childhood?				
Birth to pre-kindergarten	1	8	2	15
Pre-kindergarten	3	25	6	46
Kindergarten	4	33	3	23
Grade 1	2	17	2	15
Grade 2	1	8	2	15
Grade 3	1	8	1	8
Other	9	75	7	54

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 early childhood?				
3 years or less	0	0	2	15
4–7 years	3	25	1	8
8–11 years	1	8	4	31
12–15 years	2	17	3	23
16 years or more	6	50	3	23
Which best describes the location of your P–12 school?				
Urban	2	17	4	31
Suburban	1	8	0	0
Rural	2	17	1	8
Not currently working at the P–12 level	7	58	8	62
Considering the current and the past three school years, which content area(s) do you primarily teach? (Select all that apply)				
I do not teach in the P-12 grade levels.	7	58	8	62
Mathematics	4	33	6	46
English Language Arts & Reading	4	33	6	46
Social Studies	4	33	5	38
Science	4	33	5	38
Other	5	42	6	46
If you are college faculty, are you currently involved in the training/preparation of teacher candidates in early childhood?				
Yes	7	58	7	54
No	0	0	0	0
Not college faculty	5	42	6	46

Table D2***RLA and Social Studies (5027) Passing Score Summary by Round of Judgments***

Panelist	Panel 1		Panel 2	
	Round 1	Round 2	Round 1	Round 2
1	64.00	64.00	47.00	47.50
2	57.45	57.35	51.10	51.10
3	46.40	48.40	48.85	49.00
4	48.70	48.30	48.80	48.70
5	49.35	53.05	49.90	49.90
6	54.55	53.40	52.50	53.90
7	45.20	47.20	50.00	50.10
8	62.10	61.60	51.90	50.80
9	44.50	46.70	44.20	47.00
10	50.70	52.00	50.90	49.80
11	51.05	51.85	48.10	49.40
12	43.40	44.50	46.00	46.60
13			46.10	46.50
Average	51.45	52.36	48.87	49.25
Lowest	43.40	44.50	44.20	46.50
Highest	64.00	64.00	52.50	53.90
SD	6.79	6.03	2.51	2.08
SEJ	1.96	1.74	0.70	0.58

Table D3***Mathematics and Science (5028) Passing Score Summary by Round of Judgments***

	Panel 1		Panel 2	
Panelist	Round 1	Round 2	Round 1	Round 2
1	64.00	63.60	46.85	47.35
2	52.00	53.10	45.50	46.50
3	47.25	49.65	41.80	42.60
4	46.50	46.00	46.50	46.20
5	65.50	64.80	39.20	40.40
6	55.30	53.80	53.30	53.00
7	48.10	49.30	43.80	46.40
8	58.20	58.60	44.90	45.50
9	46.20	46.50	48.60	48.90
10	53.10	52.80	56.30	53.90
11	54.25	54.40	47.05	47.55
12	42.50	43.60	46.50	46.60
13			44.55	45.25
Average	52.74	53.01	46.53	46.93
Lowest	42.50	43.60	39.20	40.40
Highest	65.50	64.80	56.30	53.90
SD	7.18	6.68	4.45	3.62
SEJ	2.07	1.93	1.23	1.00

Table D4***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.	10	83	2	17	0	0	0	0
• The instructions and explanations provided by the facilitator were clear.	6	50	4	33	2	17	0	0
• The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	8	67	4	33	0	0	0	0
• The explanation of how the recommended passing score is computed was clear.	6	50	5	42	1	8	0	0
• The opportunity for feedback and discussion between rounds was helpful.	10	83	1	8	1	8	0	0
• The process of making the standard-setting judgments was easy to follow.	9	75	3	25	0	0	0	0

Table D4 (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 just qualified candidate	8	67	3	25	1	8		
• The between-round discussions	7	58	4	33	1	8		
• The knowledge/skills required to answer each test item	11	92	1	8	0	0		
• The passing scores of other panel members	0	0	12	100	0	0		
• My own professional experience	10	83	2	17	0	0		
Overall, how comfortable are you with the panel's recommended passing score?	Very comfortable		Somewhat comfortable		Somewhat uncomfortable		Very uncomfortable	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
• RLA and Social Studies	7	58	5	42	0	0	0	0
• Mathematics and Science	6	50	6	50	0	0	0	0
Overall, the recommended passing score is:	Too low		About right		Too high			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• RLA and Social Studies	1	8	11	92	0	0		
• Mathematics and Science	0	0	11	92	1	8		

Table D5***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.	13	100	0	0	0	0	0	0
• The instructions and explanations provided by the facilitators 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.	11	85	2	15	0	0	0	0
• The explanation of how the recommended passing score is computed was clear.	11	85	2	15	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.	12	92	1	8	0	0	0	0

Table D5 (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 just qualified candidate	11	85	2	15	0	0		
• The between-round discussions	9	69	4	31	0	0		
• The knowledge/skills required to answer each test item	12	92	1	8	0	0		
• The passing scores of other panel members	6	46	7	54	0	0		
• My own professional experience	12	92	1	8	0	0		
Overall, how comfortable are you with the panel's recommended passing score?	Very comfortable		Somewhat comfortable		Somewhat uncomfortable		Very uncomfortable	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
• RLA and Social Studies	11	85	2	15	0	0	0	0
• Mathematics and Science	13	100	0	0	0	0	0	0
Overall, the recommended passing score is:	Too low		About right		Too high			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• RLA and Social Studies	0	0	13	100	0	0		
• Mathematics and Science	0	0	13	100	0	0		



Multistate Standard-Setting Technical Report

***PRAXIS*[®] PRINCIPLES OF LEARNING AND TEACHING PREK-12 (5625)**

Licensure and Credentialing Research

ETS

Princeton, New Jersey

December 2020

EXECUTIVE SUMMARY

To support the decision-making process of education agencies in establishing a passing score (cut score) for the *Praxis*[®] Principles of Learning and Teaching (PLT): PreK-12 (5625) test, research staff from Educational Testing Service (ETS) designed and conducted a distance-based multistate standard-setting study.

PARTICIPATING STATES

Panelists from nine states and Washington, D.C., were recommended by their respective education agencies. The education agencies recommended panelists with (a) experience as either teachers or college faculty who prepare teachers and (b) familiarity with the knowledge and skills required of beginning teachers.

RECOMMENDED PASSING SCORE

ETS provides a multistate standard-setting study technical report, with the recommended passing score, to help education agencies determine an appropriate operational passing score. For the *Praxis* PLT: PreK-12 test, the recommended passing score is 55 out of a possible 85 raw-score points. The scale score associated with a raw score of 55 is 157 on a 100–200 scale.

INTRODUCTION

To support the decision-making process for education agencies establishing a passing score (cut score) for the *Praxis*[®] Principles of Learning and Teaching (PLT): PreK-12 (5625) test, research staff from Educational Testing Service (ETS) designed and conducted a distance-based multistate standard-setting study, which took place in November 2020. Education agencies¹ recommended panelists with (a) experience as either teachers or college faculty who prepare teachers and (b) familiarity with the knowledge and skills required of beginning teachers. Nine states and Washington, D.C. (Table 1) were represented by 13 panelists. (See Appendix A for the names and affiliations of the panelists.)

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

Arkansas (1 panelist)	Nevada (1 panelist)
Indiana (4 panelists)	Oregon (1 panelist)
Kansas (1 panelist)	South Dakota (1 panelist)
Kentucky (1 panelist)	Washington, D.C. (1 panelist)
Nebraska (1 panelist)	West Virginia (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 report with recommended passing score from the multistate standard-setting study to education agencies. In each state and in D.C., the department of education, the board of education, or other designated educator licensure board is responsible for establishing the operational passing score, in accordance with applicable regulations. This study provides a recommended passing score, which represents the combined judgments of a group of experienced educators. Each state, and D.C., may want to consider the recommended passing score but also other sources of information when setting the final *Praxis* PLT: PreK-12 passing score (see Geisinger & McCormick, 2010). States, or D.C., may accept the recommended passing score, adjust the score upward to reflect more stringent expectations, or adjust the score downward to reflect more lenient expectations. There is no *correct* decision; the appropriateness of any adjustment may only be evaluated in terms of its meeting each state's, or D.C.'s, needs.

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

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* PLT: PreK-12 test score and the latter, the reliability of panelists' passing-score recommendation. The SEM allows a state and D.C., to recognize that any test score on any standardized test—including a *Praxis* PLT: PreK-12 test score—is not perfectly reliable. A test score only *approximates* what a candidate truly knows or truly can do on the test. The SEM, therefore, addresses the question: How close of an approximation is the test score to the *true* score? The SEJ allows states, and D.C., to gauge the likelihood that the recommended passing score from the current panel would be similar to the passing scores recommended by other panels of experts similar in composition and experience. The smaller the SEJ, the more likely that another panel would recommend a passing score consistent with the recommended passing score. The larger the SEJ, the less likely the recommended passing score would be reproduced by another panel.

In addition to measurement error metrics (e.g., SEM, SEJ), each state, and D.C., should consider the likelihood of classification errors. That is, when adjusting a passing score, policymakers should consider whether it is more important to minimize a false-positive decision or to minimize a false-negative decision. A false-positive decision occurs when a candidate's test score suggests that he should receive a license/certificate, but his actual level of knowledge/skills indicates otherwise (i.e., the candidate does not possess the required knowledge/skills). A false-negative decision occurs when a candidate's test score suggests that she should not receive a license/certificate, but she actually possesses the required knowledge/skills. The states, and D.C., need to consider which decision error is more important to minimize.

OVERVIEW OF THE *PRAXIS*[®] PLT: PREK-12 TEST

The *Praxis*[®] PLT: PreK-12 *Test at a Glance* document (ETS, in press) describes the purpose and structure of the test. In brief, the assess' a new teacher's understanding of educational practices foundational to beginning a career as a professional educator.

The two-hour assessment contains 100 selected-response items² covering four content areas: *Students as Learners* (approximately 31 items), *Instructional Process* (approximately 30 items), *Assessment* (approximately 20 items), and *Professional Development, Leadership, and Community*

² Fifteen of the 100 multiple-choice items are pretest items and do not contribute to a candidate's score.

(approximately 19 items).³ The reporting scale for the *Praxis* PLT: PreK-12 test ranges from 100 to 200 scale-score points.

PROCESSES AND METHODS

The design of the standard-setting study included an expert panel of educators with experience with the knowledge and skills measured on the test and with new teachers or with teacher candidates. Before the study, panelists received an email explaining the purpose of the standard-setting study and requesting that they review materials for the study, such as the test specifications and an overview presentation. The review of the test specifications helped familiarize the panelists with the general structure and content of the test. The overview presentation served as an initial training in the standard-setting process its purpose.

Additionally, panelists were asked to attend a brief technology check meeting. This meeting, which lasted approximately 15 minutes, served to ensure that everyone could access the technology needed for the study. The panelists were provided with information on how to use the essential functions within the videoconferencing software. They were also guided through the process for accessing ETS's secure test review portal.

The first day of the standard-setting study began with a welcome by the meeting facilitator. After introductions of the panelists and ETS staff, the facilitator described the goals for the day. (The full agenda for the meeting is in Appendix B.) The facilitator then engaged the panel in a question and answer period about the overview presentation, followed by having the panelists review the test.

REVIEWING THE TEST

The purpose of test familiarization is for the panelists to review the test and become familiar with the manner in which a candidate would take the test. After the facilitator described the purpose of the review and how to access the test⁴, the standard-setting panelists took the test and had a discussion of the content measured. This discussion helped bring the panelists to a shared understanding of what the test measures. The test discussion covered the major content areas being addressed by the test. Panelists were asked to remark on any content areas that would be particularly challenging for entry-level teachers or

³ The number of items for each content area may vary slightly from form to form of the test.

⁴ The computer-administered test items were available through the ETS IBIS Content Review Tool.

areas that address content particularly important for entry-level teachers. Overall, this discussion serves to reduce potential judgment errors later in the standard-setting process.

DESCRIBING THE JUST QUALIFIED CANDIDATE

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

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

The final 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 typed description was made available for panelists to download and use during subsequent phases of the study (see Appendix C for the just qualified candidate description).

PANELISTS' JUDGMENTS

The standard-setting process for the *Praxis* PLT: PreK-12 test was a probability-based Modified Angoff method (Brandon, 2004; Hambleton & Pitoniak, 2006). In this study, each panelist judged each item on the likelihood (probability or chance) that the just qualified candidate would answer the item correctly. Panelists made their judgments using the following rating scale: 0, .05, .10, .20, .30, .40, .50, .60, .70, .80, .90, .95, 1. The lower the value, the less likely it is that the just qualified candidate would answer the item correctly because the item is difficult for the just qualified candidate. The higher the value, the more likely it is that the just qualified candidate would answer the item correctly.

Panelists were asked to approach the judgment process in two stages. First, they reviewed both the description of the just qualified candidate and the item and determined what was the probability that the just qualified candidate would answer the question correctly. The facilitator encouraged the panelists to consider the following rules of thumb to guide their decision:

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

Next, panelists decided how to refine their judgment within the range. For example, if a panelist thought that there was a *high chance* that the just qualified candidate would answer the question correctly, the initial decision would be in the .70 to 1 range. The second decision for the panelist was to judge if the likelihood of answering it correctly is .70, .80, .90, .95 or 1.

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

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

The panelists discussed their item-level judgments. These discussions helped panelists maintain a shared understanding of the knowledge/skills of the just qualified candidate and helped to clarify aspects of items that might not have been clear to all panelists during the Round 1 judgments. The purpose of the discussion was not to encourage panelists to conform to another's judgment, but to understand the different relevant perspectives among the panelists.

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

RESULTS

EXPERT PANELS

Table 2 presents a summary of the panelists' demographic information. The panel included 13 educators representing nine states and D.C. (See Appendix A for a listing of panelists.) Six panelists were teachers, Six were college faculty, and one held another position. Two of the three faculty members' job responsibilities included the training of teachers.

Table 2
Panel Member Demographics

	<i>N</i>	<i>%</i>
Current position		
Teacher	6	46
Administrator/Department head	0	0
College faculty	6	46
Professional Learning Coordinator/Facilitator	1	8
Race		
White	9	69
Hispanic or Latino	3	23
Asian or Asian American	1	8
Gender		
Female	8	62
Male	5	38
Are you currently certified to teach this subject in your state?		
Yes	11	85
No	2	15
Are you currently teaching this subject in your state?		
Yes	11	85
No	2	15
Are you currently supervising or mentoring other teachers of this subject?		
Yes	11	85
No	2	15
At what K–12 grade level are you currently teaching this subject?		
Elementary (K–5 or K–6)	3	23
Middle school and High school	1	8
High school (9–12 or 10–12)	3	23
Not currently teaching at the K–12 level	6	46

(continues on next page)

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

	<i>N</i>	<i>%</i>
Including this year, how many years of experience do you have teaching this subject?		
3 years or less	1	8
4–7 years	3	23
8–11 years	0	0
12–15 years	1	8
16 years or more	8	62
Which best describes the location of your K–12 school?		
Urban	1	8
Suburban	3	23
Rural	3	23
Not currently working at the K–12 level	6	46
If you are college faculty, are you currently involved in the training/preparation of teacher candidates in this subject?		
Yes	6	46
No	0	0
Not college faculty	7	54

STANDARD-SETTING JUDGMENTS

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

Table 3 also includes estimate of the measurement error associated with the judgments: the standard deviation of the mean and the standard error of judgment (SEJ). The SEJ is one way of estimating the reliability or consistency of a panel’s standard-setting judgments.⁵ It indicates how likely it would be for several other panels of educators similar in makeup, experience, and standard-setting training to the current panel to recommend the same passing score on the same form of the test.

Round 1 judgments are made without discussion among the panelists. The most variability in judgments, therefore, is typically present in the first round. Round 2 judgments, however, are informed by panel discussion; thus, it is common to see a decrease both in the standard deviation and SEJ. This decrease—indicating convergence among the panelists’ judgments—was observed (see Table 3). The Round 2 average score is the panel’s recommended passing score.

⁵ An SEJ assumes that panelists are randomly selected and that standard-setting judgments are independent. It is seldom the case that panelists are randomly sampled, and only the first round of judgments may be considered independent. The SEJ, therefore, likely underestimates the uncertainty of passing scores (Tannenbaum & Katz, 2013).

Table 3
Passing Score Summary by Round of Judgments

Panelist	Round 1	Round 2
1	48.00	48.10
2	57.70	56.10
3	59.35	57.00
4	55.00	53.60
5	56.00	55.80
6	51.20	53.10
7	57.05	56.75
8	48.30	51.50
9	58.40	58.40
10	57.60	54.80
11	42.40	47.60
12	60.85	57.00
13	61.00	59.50
Average	54.83	54.56
Lowest	42.40	47.60
Highest	61.00	59.50
SD	5.68	3.68
SEJ	1.57	1.02

The panel's passing score recommendation for the *Praxis* PLT: PreK-12 test is 54.56 (out of a possible 85 raw-score points). The value was rounded to the next highest whole number, 55, to determine the functional recommended passing score. The scale score associated with 55 raw points is 157.

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

Table 4***Passing Scores Within 1 and 2 CSEM of the Recommended Passing Score⁶***

Recommended passing score (CSEM)		Scale score equivalent
	55 (4.43)	157
-2 CSEM	47	143
-1 CSEM	51	150
+ 1 CSEM	60	165
+ 2 CSEM	64	172

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

FINAL EVALUATIONS

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

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

All but one panelist *strongly agreed* or *agreed* that they understood the purpose of the study and that the facilitator’s instructions and explanations were clear. All but one panelist *strongly agreed* or *agreed* that they were prepared to make their standard-setting judgments and that the standard-setting process was easy to follow. The panelist who selected *strongly disagree* did so for all of the questions in this set. This panelist responded positively on the training evaluation poll and the remaining questions in the final evaluation. This panelist also indicated that she was *very comfortable* with the panel’s recommendation.

All panelists reported that the description of the just qualified candidate was at least *somewhat influential* in guiding their standard-setting judgments; 10 of the 13 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 (seven of the 13 panelists) indicated that their own professional experience was *very influential* in guiding their judgments.

⁶ The unrounded CSEM value is added to or subtracted from the rounded passing-score recommendation. The resulting values are rounded up to the next-highest whole number and the rounded values are converted to scale scores.

All of the panelists indicated they were at least *somewhat comfortable* with the passing score they recommended; nine of the 13 panelists were *very comfortable*. Twelve of the 13 panelists indicated the recommended passing score was *about right* with the remaining panelist 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* PLT: PreK-12 test, research staff from ETS designed and conducted a multistate standard-setting study.

ETS provides a recommended passing score from the multistate standard-setting study to help education agencies determine an appropriate operational passing score. For the *Praxis* PLT: PreK-12 test, the recommended passing score is 55 out of a possible 85 raw-score points. The scale score associated with a raw score of 55 is 157 on a 100–200 scale.

REFERENCES

- Brandon, P. R. (2004). Conclusions about frequently studied modified Angoff standard-setting topics. *Applied Measurement in Education, 17*, 59–88.
- ETS. (in press). *The Praxis Series®: The Praxis Study Companion: Principles of Learning and Teaching: PreK-12 (5625)*. 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>
Tina Allen	Marshall University (WV)
Enid Baines	Carmel High School (IN)
Ginaelisa Cortez	Friendship Public Charter Schools (DC)
Joyce Dubach	Wawasee Community School Corporation (IN)
Michelle Fisher	Norfolk Public School (NE)
Joshua Giles	South Laurel High School (KY)
Dr. Jim Granada	Wichita State University (KS)
Liliana Jimenez	Fern Hill Elementary School (OR)
Scott Klungseth	Dakota State University (SD)
Holly Marich, Ph.D., NBCT	Northeastern Nevada Regional Professional Development Program (NV)
Benjamin Mason	Purdue University (IN)
Clinton Petty	Southern Arkansas University (AR)
Youjin Yang	Saint Mary-of-the-Woods College (IN)

APPENDIX B
STUDY AGENDA

AGENDA

***Praxis*[®] Principles of Learning and Teaching PreK-12 (5625) Standard-Setting Study**

Day 1

Welcome and Introduction

Overview of Standard Setting and the *Praxis* Principles of Learning and Teaching (PLT): PreK-12 Test

Review the *Praxis* PLT: PreK-12 Test

Discuss the *Praxis* PLT: PreK-12 Test

Lunch

Overview of the Just Qualified Candidate

Define the Just Qualified Candidate

Break

Define the Just Qualified Candidate (continued)

End of Day 1

AGENDA

***Praxis*[®] Principles of Learning and Teaching PreK-12 (5625) Standard-Setting Study**

Day 2

Overview of Day 2

Review Just-Qualified Candidate definition

Standard-setting training presentation

Practice Round: Selected-response standard-setting judgments

Break

Practice Round: Data Discussion

Evaluation of the standard-setting training

Lunch

Round 1: Selected-response standard-setting judgments

Break

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

End of Day 2

AGENDA

***Praxis*[®] Principles of Learning and Teaching PreK-12 (5625) Standard-Setting Study**

Day 3

Overview of Day 3

Round 1 Feedback and Round 2 Judgments

Break

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

Break

Feedback on Round 2 Recommended Cut Score

Complete Final Evaluation

End of Study

APPENDIX C

JUST QUALIFIED CANDIDATE DESCRIPTION

Description of the Just Qualified Candidate⁷

A just qualified candidate ...

1. Understands how basic concepts of learning theory, motivational theory, and human development impact the instructional process
2. Knows some ways in which major learner variables—such as gender, language, culture, socioeconomic status, background knowledge, and areas of exceptionality (e.g., students with disabilities, gifted students) affect student learning and performance
3. Knows a variety of developmentally appropriate instructional strategies and resources for supporting student learning (e.g., differentiation, modeling, small groups, learning centers, modalities of presentation, etc.) and to elicit a variety of student responses (e.g., verbal, non-verbal, technology based)
4. Knows the principles and strategies for creating an inclusive and positive learning environment through classroom management and procedures
5. Knows the components of effective questioning, listening strategies, and communication techniques and how they are used in different contexts to foster a supportive and effective learning environment
6. Knows how to create and select instructional objectives and aligns appropriate instructional strategies, assessments, and resources to achieve them
7. Understands the types and uses of assessments (e.g. formative, summative, normed, criterion-referenced); can select appropriate assessments, interpret the results, and use the results to guide instruction
8. Knows the legal responsibilities of teachers and the implications of major educational legislation and court decisions
9. Is familiar with a variety of professional development practices and resources (e.g., learning communities, professional literature and associations, workshops and conferences, self-reflection)
10. Understands the basics of inter-disciplinary teamwork and collaboration with stakeholders

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

APPENDIX D

FINAL EVALUATION RESULTS

Table D1***Final Evaluation***

	Strongly agree		Agree		Disagree		Strongly disagree	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
• I understood the purpose of this study.	11	85	1	8	0	0	1	8
• The instructions and explanations provided by the facilitators were clear.	11	85	1	8	0	0	1	8
• The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	10	77	2	15	0	0	1	8
• The explanation of how the recommended passing score is computed was clear.	11	85	1	8	0	0	1	8
• The opportunity for feedback and discussion between rounds was helpful.	10	77	2	15	0	0	1	8
• The process of making the standard-setting judgments was easy to follow.	10	77	2	15	0	0	1	8

Table D1 (continued)

Final Evaluation

How influential was each of the following factors in guiding your standard-setting judgments?	Very influential		Somewhat influential		Not influential			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• The description of the just qualified candidate	10	77	3	23	0	0		
• The between-round discussions	9	69	4	31	0	0		
• The knowledge/skills required to answer each test item	11	85	2	15	0	0		
• The passing scores of other panel members	2	15	11	85	0	0		
• My own professional experience	7	54	6	46	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	69	4	31	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:	1	8	12	92	0	0		



Multistate Standard-Setting Technical Report

***PRAXIS*[®] ELEMENTARY EDUCATION ASSESSMENT (5006):
READING AND LANGUAGE ARTS AND SOCIAL STUDIES SUBTEST
(5007)
MATHEMATICS AND SCIENCE SUBTEST (5008)**

Licensure and Credentialing Research

ETS

Princeton, New Jersey

January 2021

EXECUTIVE SUMMARY

To support the decision-making process of education agencies establishing a passing score (cut score) for the *Praxis*[®] Elementary Education Assessment (5006), research staff from Educational Testing Service (ETS) designed and conducted a distance-based multistate standard-setting study.

PARTICIPATING STATES

Panelists from six states and Washington, D.C. were recommended by their respective education agencies. The education agencies recommended panelists with (a) experience as either elementary teachers or college faculty who prepare elementary teachers and (b) familiarity with the knowledge and skills required of beginning elementary teachers.

RECOMMENDED PASSING SCORE

ETS provides a multistate standard-setting study technical report, with the recommended passing score, to help education agencies determine appropriate operational passing scores for the two subtests of the Praxis Elementary Education Assessment. The panelists completed the standard-setting process and recommended two scores, one for each subtest—Reading and Language Arts and Social Studies (5007) and Mathematics and Science (5008). The recommended passing scores are listed in Table 1. The scale scores associated with each of the raw score recommendations are on a 100–200 scale.

Table 1.
Recommended Passing Scores (raw values and scale scores)

Subtest Title	Test Code	Recommended raw score	Total raw scores	Praxis scale score
Reading and Language Arts and Social Studies	5007	53	80	160
Mathematics and Science	5008	45	75	158

INTRODUCTION

To support the decision-making process for education agencies establishing passing scores (cut scores) for the *Praxis*[®] Elementary Education Assessment (5006), research staff from Educational Testing Service designed and conducted a distance-based multistate standard-setting study in December 2020. Education agencies¹ recommended panelists with (a) experience as either elementary teachers or college faculty who prepare elementary teachers and (b) familiarity with the knowledge and skills required of beginning elementary teachers. Six states and Washington, D.C. (Table 2) were represented by 18 panelists. (See Appendix A for the names and affiliations of the panelists.)

Table 2.
Participating States, Washington, D.C., and Number of Panelists

Alaska (2 panelists)	Nevada (1 panelist)
Arkansas (3 panelists)	Washington, D.C. (2 panelists)
Idaho (3 panelists)	West Virginia (3 panelists)
Indiana (4 panelists)	

The following technical report contains three sections. The first section describes the content and format of the test. The second section describes the standard-setting processes and methods. The third section presents the results of the standard-setting study.

ETS provides a recommended passing score from the multistate standard-setting study to education agencies. In each state and in D.C., the department of education, the board of education, or a designated educator licensure board is responsible for establishing the operational passing score in accordance with applicable regulations. This study provides a recommended passing score,² which represents the combined judgments of two panels of experienced educators. Each state and D.C. may want to consider the recommended passing score but also other sources of information when setting the final *Praxis* Elementary Education Assessment passing score (see Geisinger & McCormick, 2010). A state, or D.C., may accept the recommended passing score, adjust the score upward to reflect more stringent expectations, or adjust the score downward to reflect more lenient expectations. There is no *correct* decision; the appropriateness of any adjustment may only be evaluated in terms of its meeting the state and D.C.'s needs.

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

² In addition to the recommended passing score averaged across the two panels, the recommended passing scores for each panel are presented.

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* Elementary Education Assessment scores and the latter, the reliability of panelists' passing-score recommendations. The SEM allows a state and D.C. to recognize that any test score on any standardized test—including a *Praxis* Elementary Education Assessment score—is not perfectly reliable. A test score only *approximates* what a candidate truly knows or truly can do on the test. The SEM, therefore, addresses the question: How close of an approximation is the test score to the *true* score? The SEJ allows a state and D.C. to gauge the likelihood that the recommended passing score from a particular panel would be similar to the passing scores recommended by other panels of experts similar in composition and experience. The smaller the SEJ, the more likely that another panel would recommend a passing score consistent with the recommended passing score. The larger the SEJ, the less likely the recommended passing score would be reproduced by another panel.

In addition to measurement error metrics (e.g., SEM, SEJ), each state and D.C., should consider the likelihood of classification errors. That is, when adjusting a passing score, policymakers should consider whether it is more important to minimize a false-positive decision or to minimize a false-negative decision. A false-positive decision occurs when a candidate's test score suggests that he should receive a license/certificate, but his actual level of knowledge/skills indicates otherwise (i.e., the candidate does not possess the required knowledge/skills). A false-negative decision occurs when a candidate's test score suggests that she should not receive a license/certificate, but she actually does possess the required knowledge/skills. The state, or D.C., needs to consider which decision error is more important to minimize.

OVERVIEW OF THE *PRAXIS*[®] ELEMENTARY EDUCATION ASSESSMENT

The *Praxis*[®] Elementary Education *Study Companion* document (ETS, in press) describes the purpose and structure of the test. In brief, the test measures whether entry-level elementary teachers have the knowledge/skills believed necessary for competent professional practice.

Candidates must pass both subtests in order to pass the *Praxis* Elementary Education Assessment. They may retake the subtest that they did not pass in order to fulfill the requirement. The reporting scale for each subtest ranges from 100 to 200.

READING AND LANGUAGE ARTS AND SOCIAL STUDIES (5007)

This 2-hour subtest contains 95 selected-response items covering two content areas: *Reading and Language Arts* (approximately 62 items) and *Social Studies* (approximately 33 items).³ The reporting scale for the *Praxis* Reading and Language Arts and Social Studies subtest ranges from 100 to 200 scale-score points.

MATHEMATICS AND SCIENCE (5008)

This 2-hour subtest contains 85 selected-response items covering two content areas: *Mathematics* (approximately 51 items) and *Science* (approximately 34 items).³ The reporting scale for the *Praxis* Mathematics and Science subtest ranges from 100 to 200 scale-score points.

PROCESSES AND METHODS

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

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

REVIEWING THE TEST

Test familiarization was the first activity for the panel. The purpose of test familiarization is for the panelists to review the test and become familiar with the manner in which a candidate would take the test. After the facilitator described the purpose of the review and how to access the test⁴, the standard-

³ The number of items for each content area may vary slightly from form to form of the test.

⁴ The computer-administered test items were available through the ETS IBIS Content Review Tool.

setting panelists took the test and had a discussion of the content measured. This discussion helped bring the panelists to a shared understanding of what the test measures.

The test discussion covered the major content areas being addressed by the test. Panelists were asked to remark on any content areas that would be particularly challenging for entry-level teachers or areas that address content particularly important for entry-level teachers. Overall, this discussion serves to reduce potential judgment errors later in the standard-setting process.

For this two-panel, multistate standard-setting study, the first test that was reviewed differed for each panel. The educators on Panel 1 first viewed the Reading and Language Arts (RLA) and Social Studies subtest. The educators on Panel 2 reviewed the Mathematics (Math) and Science subtest as their first activity. The facilitator took notes of each panel's content discussion. After completing the entire standard-setting process on the first subtest, each panel would then switch to the other subtest. The notes from the content discussion were provided to help each panel see the comparisons in the discussion points.

DESCRIBING THE JUST QUALIFIED CANDIDATE

Following the review of the test, panelists worked together to describe the just qualified candidate—specific to the content measured on the subtest the panel reviewed. The *just qualified candidate description* plays a central role in standard setting (Perie, 2008); the goal of the standard-setting process is to identify the test score that aligns with this description.

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

The final 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 typed description was made available for panelists to download and use during subsequent phases of the study (see Appendix C for the just qualified candidate descriptions).

Panel 1 created the just qualified candidate description, specific to the knowledge and skills measured on the RLA and Social Studies subtest. Panel 2 created the just qualified candidate

description for the Math and Science subtest. Each panel completed their test familiarization for the second subtest, aided by the discussion notes from the other panel. Afterwards, they worked to understand the just qualified candidate description of that second subtest. So, Panel 1 received the description created by Panel 2 for the Math and Science subtest; Panel 2 received the description created by Panel 1 for the RLA and Social Studies subtest.

To *understand* the just qualified candidate description provided by the other panel, each set of panelists worked together to create performance indicators—observable, or measurable, actions that a candidate could do to demonstrate the level of knowledge and skills described in the description. The panel would discuss these statements, specifying what the just qualified candidate *could do*, as well as what the candidate *could not do*.

PANELISTS' JUDGMENTS

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

Panelists were asked to approach the judgment process in two stages. First, they reviewed both the description of the just qualified candidate and the item and determined what was the probability that the just qualified candidate would answer the question correctly. The facilitator encouraged the panelists to consider the following rules of thumb to guide their decision:

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

Next, panelists decided how to refine their judgment within the range. For example, if a panelist thought that there was a high chance that the just qualified candidate would answer the question correctly, the initial decision would be in the .70 to 1 range. The second decision for the panelist was to judge if the likelihood of answering it correctly is .70, .80, .90, .95 or 1.

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

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

The panelists discussed their item-level judgments. These discussions helped panelists maintain a shared understanding of the knowledge/skills of the just qualified candidate and helped to clarify aspects of items that might not have been clear to all panelists during the Round 1 judgments. The purpose of the discussion was not to encourage panelists to conform to another's judgment, but to understand the different relevant perspectives among the panelists.

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

Each panel completed both rounds of judgments on the first subtest before starting over with the test familiarization for the second subtest. After understanding the just qualified candidate description for their second subtest, each panel completed both rounds of judgments for their second subtest. Other than content discussion notes and the description of the just qualified candidate, results from Panel 1 were not shared with Panel 2. The judgments, results, and item-level discussions for Panel 2 were independent of judgments, results, and item-level discussions that occurred with Panel 1.

RESULTS

EXPERT PANELS

Table 3 presents a summary of the panelists' demographic information. The panel included 18 educators representing six states and D.C. (See Appendix A for a listing of panelists.) Two panelists were teachers and 16 were college faculty. All of the faculty members' job responsibilities included the training of elementary teachers. The number of experts by panel and their demographic information are presented in Appendix D (Table D1).

Table 3.
Panel Member Demographics (Across Panels)

	<i>N</i>	<i>%</i>
Current position		
Teacher	2	11
Administrator/Department head	0	0
College faculty	16	89
Other	0	0
Race		
White	17	94
Asian or Asian American	1	6
Gender		
Female	15	83
Male	3	17
Are you currently certified to teach elementary education in your state?		
Yes	7	39
No	11	61
Are you currently teaching Elementary Education in your state?		
Yes	6	33
No	12	67
Are you currently supervising or mentoring other teachers of elementary education?		
Yes	14	78
No	4	22

(continues on next page)

Table 3. (continued)
Panel Member Demographics (Across Panels)

	<i>N</i>	<i>%</i>
At what P–12 grade level are you currently teaching elementary education? (Select all that apply)		
Grade 1	0	0
Grade 2	0	0
Grade 3	1	6
Grade 4	1	6
Grade 5	0	0
Other	16	89
Including this year, how many years of experience do you have teaching elementary education?		
3 years or less	4	22
4–7 years	3	17
8–11 years	2	11
12–15 years	3	17
16 years or more	6	33
Which best describes the location of your P–12 school?		
Urban	0	0
Suburban	1	6
Rural	1	6
Not currently working at the P–12 level	16	89
If you are college faculty, are you currently involved in the training/preparation of teacher candidates in this subject?		
Yes	16	89
No	0	0
Not college faculty	2	11

STANDARD-SETTING JUDGMENTS

Tables 4 and 5 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.⁵ It indicates how likely it would be for several other panels of educators similar in makeup, experience, and standard-setting training to the current panel to

⁵ An SEJ assumes that panelists are randomly selected and that standard-setting judgments are independent. It is seldom the case that panelists are randomly sampled, and only the first round of judgments may be considered independent. The SEJ, therefore, likely underestimates the uncertainty of passing scores (Tannenbaum & Katz, 2013).

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 (Tables D2 and D3).

Table 4.
Reading and Language Arts and Social Studies Summary of Round 2 Standard-setting Judgments Per Panel

Test Code		Panel 1	Panel 2
5007	Average	53.02	52.81
	Lowest	45.20	47.80
	Highest	61.75	60.70
	SD	4.71	4.26
	SEJ	1.57	1.42

Table 5.
Mathematics and Science Summary of Round 2 Standard-setting Judgments Per Panel

Test Code		Panel 1	Panel 2
5008	Average	45.22	44.33
	Lowest	41.90	41.15
	Highest	48.70	50.85
	SD	1.94	2.91
	SEJ	0.65	0.97

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* Elementary Education Assessment are listed in Table 6. The values were rounded to the next highest whole number, to determine the functional recommended passing scores. There are 80 raw score points available for the Reading and Language Arts and Social Studies subtest (5007). For the Mathematics and Science subtest (5008) there are 70 available raw score points.

Table 6.
Recommended Passing Scores per Panel and Subtest

Subtest: Panel	Recommended Score		Praxis Score (100 – 200 Scale)
	Raw	Rounded	
RLA & Social Studies 5007: Panel 1	53.02	54	162
RLA & Social Studies 5007: Panel 2	52.81	53	160
Math & Science 5008: Panel 1	45.22	46	160
Math & Science 5008: Panel 2	44.33	45	158

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. Table 7 displays those results.

Table 7.
Recommended Passing Scores per Subtest (Panels Scores Combined)

Subtest	Panels' Averaged Recommendations		Praxis Score (100 – 200 Scale)
	Raw	Rounded	
RLA & Social Studies 5007	52.92	53	160
Math & Science 5008	44.78	45	158

Tables 8 and 9 present the estimated conditional standard error of measurement (CSEM) around the recommended passing score. A standard error represents the uncertainty associated with a test score. The scale scores associated with one and two CSEM above and below the recommended passing score are provided. The conditional standard error of measurement provided is an estimate.

Table 8.

Reading and Language Arts and Social Studies Passing Scores Within 1 and 2 CSEM of the Recommended Passing Score⁶

	Recommended passing score (CSEM)		Scale score equivalent
		53 (4.26)	160
Test: 5007	-2 CSEM	45	146
	-1 CSEM	49	153
	+ 1 CSEM	58	169
	+ 2 CSEM	62	176

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

Table 9.

Mathematics and Science Passing Scores Within 1 and 2 CSEM of the Recommended Passing Score⁶

	Recommended passing score (CSEM)		Scale score equivalent
		45 (4.04)	158
Test: 5008	-2 CSEM	37	143
	-1 CSEM	41	151
	+ 1 CSEM	50	168
	+ 2 CSEM	54	176

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

FINAL EVALUATIONS

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

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

All panelists *strongly agreed* or *agreed* that they understood the purpose of the study 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* that the standard-setting process was easy to follow.

⁶ The unrounded CSEM value is added to or subtracted from the rounded passing-score recommendation. The resulting values are rounded up to the next-highest whole number and the rounded values are converted to scale scores.

All panelists reported that the description of the just qualified candidate was at least *somewhat influential* in guiding their standard-setting judgments; 16 of the 18 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. Seventeen of the 18 panelists indicated that their own professional experience was *very influential* in guiding their judgments.

For the Reading and Language Arts and Social Studies subtest, all of the panelists indicated they were at least *somewhat comfortable* with the passing score they recommended; 17 of the 18 panelists were *very comfortable*. All of the 25 panelists indicated the recommended passing score was *about right*.

For the Mathematics and Science subtest, all of the panelists indicated they were at least *somewhat comfortable* with the passing score they recommended; 17 of the 18 panelists were *very comfortable*. Seventeen of the 18 panelists indicated the recommended passing score was *about right* with the remaining panelist 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* Elementary Education Assessment, research staff from ETS designed and conducted a distance-based multistate standard-setting study.

ETS provides a recommended passing score from the multistate standard-setting study to help education agencies determine appropriate operational passing scores. The panelists from each of the two panels completed the standard-setting process and recommended⁷ two scores—one for each of the two subjects. The recommended passing scores are listed in Table 10. The scale scores listed associated with each of the raw score recommendations are on a 100–200 scale.

Table 10.
Recommended Passing Scores (raw values and scale scores)

Subtest Title	Test Code	Recommended raw score	Total raw scores	Praxis scale score
Reading and Language Arts and Social Studies	5007	53	80	160
Mathematics and Science	5008	45	75	158

⁷ Results from the two panels participating in the study were averaged to produce the recommended passing score.

REFERENCES

- Brandon, P. R. (2004). Conclusions about frequently studied modified Angoff standard-setting topics. *Applied Measurement in Education, 17*, 59–88.
- ETS. (in press). *The Praxis Series®: The Praxis Study Companion: Elementary Education (5006)*. 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>
Mary Carlson	University of Alaska (AK)
Angela Curfman	West Liberty University (WV)
Kitty Deal	University of Alaska Southeast (AK)
Jennifer Gardner	Lewis-Clark State College (ID)
Brian Herndon	John Brown University (AR)
Marsha Moore	American College of Education (IN)
Denise Padilla	Great Basin College (NV)
Matthew Perkins Coppola	Purdue University Fort Wayne (IN)
Kirsten Pomerantz	University of Idaho, Coeur d'Alene (ID)
Renee Powell	Harrison County Schools (WV)
Suad Sakalli Gumus	Saint Mary of the Woods College (IN)
Deanna Santoro	Relay Graduate School of Education (DC)
Krystal Shipp	Arkansas Tech University (AR)
Mark Shore	University of the District of Columbia (DC)
Stefanie Sorbet	University of Central Arkansas (AR)
Melinda Tompkins	Lewis Clark State College (ID)
April Wallace	Kellogg Elementary School (WV)
Ilfa Zhulamanova	University of Southern Indiana (IN)

APPENDIX B
STUDY AGENDA

AGENDA

***Praxis*[®] Elementary Education Assessment (5006) Standard-Setting Study**

Day 1

Welcome and Introduction

Standard-setting overview presentation

- Q&A about the training

Test familiarization for the first subtest

- Panel 1: Reading & Language Arts (RLA) & Social Studies (5007) subtest
- Panel 2: Mathematics & Science (5008) subtest

Break

Discuss the content measured on the first subtest

Lunch Break

Overview of the Just Qualified Candidate

Define the Just Qualified Candidate for the subtest

- Panel 1: Reading & Language Arts (RLA) & Social Studies (5007) subtest
- Panel 2: Mathematics & Science (5008) subtest

Break

Define the Just Qualified Candidate for the subtest (continued)

Standard-setting training presentation

Practice Round: Selected-response standard-setting judgments

End of Day 1

AGENDA

***Praxis*[®] Elementary Education (5006) Standard-Setting Study**

Day 2

Overview of the Day

Practice Round: Data Discussion

Evaluation of the standard-setting training (poll)

Round 1: Selected-response standard-setting judgments

Break

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

Lunch Break

Round 1 Feedback Summary data

Evaluation of the data presentation (poll)

Round 1 feedback: Item-level data and Round 2 judgments

Break

Round 1 feedback: Item-level data and Round 2 judgments (*continued*)

Test familiarization for the second subtest

- Panel 1: Mathematics & Science (5008)
- Panel 2: Reading Language Arts (RLA) & Social Studies (5007)

Wrap up and Complete Day 2

AGENDA

***Praxis*[®] Elementary Education (5006) Standard-Setting Study**

Day 3

Overview of the Day

Discuss the content measured on the second subtest

Reviewing the Just-Qualified Candidate – Whole group discussion

- Panel 1: Mathematics & Science (5008) subtest
- Panel 2: Reading & Language Arts (RLA) & Social Studies (5007) subtest

Break

Round 1: Selected-response standard-setting judgments

Lunch Break

Round 1 Feedback Summary data

Evaluation of the data presentation (poll)

Round 1 Feedback and Round 2 Judgments

Break

Feedback on Round 2 Recommended Cut Scores

Complete Final Evaluation

End of Study

APPENDIX C

JUST QUALIFIED CANDIDATE DESCRIPTION

Description of the Just Qualified Candidate⁸

A just qualified candidate ...

Reading and Language Arts

1. Understands best practices of teaching the five components (Phonics, phonemic awareness, fluency, comprehension, vocabulary) of reading literacy.
2. Knows the characteristics of effective writing (applying organization of ideas, develop main ideas and supporting details, varied sentence structure, word choice, mechanics) and how to give developmentally and age-appropriate feedback.
3. Understands the relationship of text structures and features to comprehension development.
4. Understands and applies conventions of grammar, usage, mechanics and spelling when reading and writing.
5. Understands how to review student work and make recommendations based on identified strengths and weaknesses to inform instruction.
6. Is familiar with and can identify figurative language in reading and writing.

Social Studies

7. Is familiar with the effects of historical events on economics, government and society.
8. Understands how to help students use primary resources and research materials in social studies
9. Knows how people of different cultural backgrounds interact with their environment, self, family, neighborhood and organizations.
10. Knows general themes of geography (place, location, movement, interaction, regions) and is familiar with its impact on physical and human systems.
11. Knows how to use literacy strategies to support social studies learning.
12. Knows the basic structure of the United States Government.
13. Knows characteristics of different systems of government.

⁸ Description of the just qualified candidate focuses on the knowledge/skills that differentiate a *just* from a *not quite* qualified candidate.

Description of the Just Qualified Candidate (continued)

A just qualified candidate ...

Mathematics

1. Understands basic and precise mathematical terminology (e.g., mode, median, distributive and commutative properties); defines and correctly applies the terminology to real world and theoretical applications.
2. Demonstrates knowledge of critical mathematical concepts from the elementary curriculum in the four categories (numbers & operations; algebraic thinking; geometry & measurement; data, statistics and probability) by using the knowledge
 - a. to solve problems,
 - b. to identify examples of concepts,
 - c. to do student error analysis,
 - d. to differentiate instruction on mathematical concepts, and
 - e. to explain the validity and logic of results

Science

3. Knows principles of scientific inquiry—e.g., the scientific method-- and can apply to them to teaching specific scientific concepts, including making connections to the real world.
4. Applies common safety procedures when engaging in scientific inquiry and utilizes the safety procedures when implementing lessons, including hands-on activities and experiments.
5. Knows the ways in which humans impact the environment (e.g. erosion, climate change) and demonstrates the ways in which actions align to those impacts (e.g. planting trees, recycling).
6. Comprehends and applies basic physical science concepts (energy and matter) and their interactions within authentic life experiences.
7. Comprehends and applies basic life science concepts (plants and animals) and their interactions within authentic experiences both inside and outside the classroom.
8. Understands the relationship between celestial bodies and can show/illustrate how they interact (e.g. revolution, rotation, orbit) and are viewed from earth.
9. Understands the relationship among earth systems and materials (e.g., the water cycle, climate)
10. and can show/illustrate how they interact and how systems change over time
11. Defines and integrates elementary scientific terminology consistently and accurately in the context of appropriate grade-level instruction.
12. Collects and interprets scientific data through visual representations (e.g. tables, graphs)

Description of the Just Qualified Candidate (continued)

A just qualified candidate ...

13. Recognizes patterns of errors in student work that indicates misconceptions in foundational knowledge.

APPENDIX D

RESULTS

Table D1
Panel Member Demographics (by Panel)

	Panel 1		Panel 2	
	<i>N</i>	%	<i>N</i>	%
Current position				
Teacher	1	11	1	11
College faculty	8	89	8	89
Race				
White	8	89	9	100
Asian or Asian American	1	11	0	0
Gender				
Female	7	78	8	89
Male	2	22	1	11
Are you currently certified to teach elementary education in your state?				
Yes	4	44	3	33
No	5	56	6	67
Are you currently teaching elementary education in your state?				
Yes	3	33	3	33
No	6	67	6	67
Are you currently supervising or mentoring other teachers of elementary education?				
Yes	6	67	8	89
No	3	33	1	11
At what P-12 grade level are you currently teaching elementary education?				
Grade 1	0	0	0	0
Grade 2	0	0	0	0
Grade 3	1	11	0	0
Grade 4	0	0	1	11
Grade 5	0	0	0	0
Other	8	89	8	89

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 elementary education?				
3 years or less	3	33	1	11
4–7 years	1	11	2	22
8–11 years	0	0	2	22
12–15 years	1	11	2	22
16 years or more	4	44	2	22
Which best describes the location of your P–12 school?				
Urban	0	0	0	0
Suburban	1	11	0	0
Rural	0	0	1	11
Not currently working at the P–12 level	8	89	8	89
If you are college faculty, are you currently involved in the training/preparation of teacher candidates in this subject?				
Yes	8	89	8	89
No	0	0	0	0
Not college faculty	1	11	1	11

Table D2***RLA and Social Studies (5007) Passing Score Summary by Round of Judgments***

Panelist	Panel 1		Panel 2	
	Round 1	Round 2	Round 1	Round 2
1	60.20	57.90	46.10	47.80
2	53.00	52.95	51.60	51.70
3	52.10	53.20	54.90	55.80
4	65.15	61.75	57.90	56.00
5	48.95	50.70	61.50	60.70
6	54.00	53.90	50.20	47.90
7	50.30	50.90	50.50	50.40
8	41.00	45.20	50.30	50.60
9	47.60	50.70	54.65	54.35
Average	52.48	53.02	53.07	52.81
Lowest	41.00	45.20	46.10	47.80
Highest	65.15	61.75	61.50	60.70
SD	7.03	4.71	4.66	4.26
SEJ	2.34	1.57	1.55	1.42

Table D3***Mathematics and Science (5008) Passing Score Summary by Round of Judgments***

Panelist	Panel 1		Panel 2	
	Round 1	Round 2	Round 1	Round 2
1	46.60	46.50	42.40	42.90
2	49.30	48.70	45.55	44.75
3	43.80	44.60	40.40	43.20
4	39.90	41.90	46.65	46.55
5	43.20	43.60	44.70	44.50
6	46.80	45.90	45.20	41.15
7	43.10	44.50	42.45	42.95
8	45.00	45.00	42.20	42.10
9	44.80	46.30	53.40	50.85
Average	44.72	45.22	44.77	44.33
Lowest	39.90	41.90	40.40	41.15
Highest	49.30	48.70	53.40	50.85
SD	2.69	1.94	3.80	2.91
SEJ	0.90	0.65	1.27	0.97

Table D4***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.	9	100	0	0	0	0	0	0
• The instructions and explanations provided by the facilitators were clear.	7	78	2	22	0	0	0	0
• The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	8	89	1	11	0	0	0	0
• The explanation of how the recommended passing score is computed was clear.	8	89	1	11	0	0	0	0
• The opportunity for feedback and discussion between rounds was helpful.	9	100	0	0	0	0	0	0
• The process of making the standard-setting judgments was easy to follow.	9	100	0	0	0	0	0	0

Table D4 (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 just qualified candidate	9	100	0	0	0	0		
• The between-round discussions	8	89	1	11	0	0		
• The knowledge/skills required to answer each test item	8	89	1	11	0	0		
• The passing scores of other panel members	5	56	4	44	0	0		
• My own professional experience	8	89	1	11	0	0		
Overall, how comfortable are you with the panel's recommended passing score?	Very comfortable		Somewhat comfortable		Somewhat uncomfortable		Very uncomfortable	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
• RLA and Social Studies	8	89	1	11	0	0	0	0
• Mathematics and Science	9	100	0	0	0	0	0	0
Overall, the recommended passing score is:	Too low		About right		Too high			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• RLA and Social Studies	0	0	9	100	0	0		
• Mathematics and Science	0	0	9	100	0	0		

Table D5***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.	8	89	1	11	0	0	0	0
• The instructions and explanations provided by the facilitators were clear.	7	78	2	22	0	0	0	0
• The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	8	89	1	11	0	0	0	0
• The explanation of how the recommended passing score is computed was clear.	9	100	0	0	0	0	0	0
• The opportunity for feedback and discussion between rounds was helpful.	9	100	0	0	0	0	0	0
• The process of making the standard-setting judgments was easy to follow.	9	100	0	0	0	0	0	0

Table D5 (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 just qualified candidate	7	78	2	22	0	0		
• The between-round discussions	7	78	2	22	0	0		
• The knowledge/skills required to answer each test item	8	89	1	11	0	0		
• The passing scores of other panel members	5	56	3	33	1	11		
• My own professional experience	9	100	0	0	0	0		
Overall, how comfortable are you with the panel's recommended passing score?	Very comfortable		Somewhat comfortable		Somewhat uncomfortable		Very uncomfortable	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
• RLA and Social Studies	9	100	0	0	0	0	0	0
• Mathematics and Science	8	89	1	11	0	0	0	0
Overall, the recommended passing score is:	Too low		About right		Too high			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• RLA and Social Studies	0	0	9	100	0	0		
• Mathematics and Science	1	11	8	89	0	0		