

## Multistate Standard-Setting Technical Report

# **PRAXIS® CORE ACADEMIC SKILLS FOR EDUCATORS: MATHEMATICS (5733)**

Student and Teacher Assessments: Validity and Test Use

ETS

Princeton, New Jersey

March, 2019

---

# EXECUTIVE SUMMARY

The the *Praxis*<sup>®</sup> Core Academic Skills for Educators assessment consists of three subtests (Reading, Writing, and Mathematics). The Mathematics (5733) subtest has been revised to reflect new standards<sup>1</sup>. To support the decision-making process of education agencies establishing a passing score (cut score) for the *Praxis* Core Academic Skills for Educators: Mathematics (5733) subtest, research staff from Educational Testing Service (ETS) designed and conducted a multistate standard-setting study.

## PARTICIPATING STATES

Panelists from 20 states and Washington, DC were recommended by their respective education agencies. The education agencies recommended panelists with (a) experience preparing teacher candidates and (b) familiarity with the knowledge and skills required of candidates entering a teacher preparation program.

## RECOMMENDED PASSING SCORE

ETS provides a recommended passing score from the multistate standard-setting study to help education agencies determine an appropriate operational passing score. For the *Praxis* Core Academic Skills for Educators: Mathematics subtest, the recommended passing score<sup>2</sup> is 28 out of a possible 50 raw-score points. The scale score associated with a raw score of 28 is 150 on a 100–200 scale.

For the remaining two subtests, ETS conducted a multistate standard-setting study in February 2013. The recommended passing scores of 156 for the Reading subtest and 162 for the Writing subtest. This information is also available on the [ETS website](#).

---

<sup>1</sup> The Reading (5712) and Writing (5722) subtests were reviewed by educator preparation faculty and a national advisory committee. It was determined that the content domains did not need revision.

<sup>2</sup> Results from the two panels participating in the study were averaged to produce the recommended passing score.

The the *Praxis*<sup>®</sup> Core Academic Skills for Educators assessment consists of three subtests (Reading, Writing, and Mathematics). The Mathematics (5733) subtest has been revised to reflect new standards<sup>3</sup>. To support the decision-making process for education agencies establishing a passing score (cut score) for the *Praxis*<sup>®</sup> Core Academic Skills for Educators: Mathematics (5733) subtest, research staff from ETS designed and conducted a multistate standard-setting study in February 2019 in Princeton, New Jersey. Education agencies<sup>4</sup> recommended panelists with (a) experience preparing teachers candidates and (b) familiarity with the knowledge and skills required of candidates entering a teacher preparation program. Twenty states and Washington, DC (Table 1) were represented by 33 panelists. (See Appendix A for the names and affiliations of the panelists.)

**Table 1**  
***Participating Jurisdictions and Number of Panelists***

---

Alaska (1 panelist)	Maryland (2 panelists)
Alabama (2 panelists)	Mississippi (2 panelists)
Arkansas (2 panelists)	Nebraska (1 panelist)
Connecticut (1 panelist)	New Jersey (3 panelists)
District of Columbia (1 panelist)	Nevada (1 panelist)
Georgia (3 panelists)	North Carolina (2 panelists)
Hawaii (1 panelist)	Pennsylvania (2 panelists)
Idaho (1 panelist)	South Carolina (2 panelists)
Iowa (1 panelist)	Tennessee (1 panelist)
Kansas (1 panelist)	West Virginia (2 panelists)
Louisiana (1 panelist)	

---

The following technical report contains three sections. The first section describes the content and format of the subtest. Although the *Praxis* Core Academic Skills for Educators assessment consists of three subtests, the description in this report will focus solely on the Mathematics (5733) subtest. The second section describes the standard-setting processes and methods. The third section presents the results of the standard-setting study.

ETS provides a recommended passing score from the multistate standard-setting study to education agencies. In each jurisdiction, the department of education, the board of education, or a

---

<sup>3</sup> The Reading (5712) and Writing (5722) subtests were reviewed by educator preparation faculty and a national advisory committee. It was determined that the content domains did not need revision.

<sup>4</sup> States and jurisdictions that currently use *Praxis* tests were invited to participate in the multistate standard-setting study.

designated educator licensure board is responsible for establishing the operational passing score in accordance with applicable regulations. This study provides a recommended passing score,<sup>5</sup> which represents the combined judgments of two panels of experienced educators. Each jurisdiction may want to consider the recommended passing score but also other sources of information when setting the final *Praxis* Core Academic Skills for Educators: Mathematics passing score (see Geisinger & McCormick, 2010). A jurisdiction may accept the recommended passing score, adjust the score upward to reflect more stringent expectations, or adjust the score downward to reflect more lenient expectations. There is no *correct* decision; the appropriateness of any adjustment may only be evaluated in terms of its meeting the jurisdiction's needs.

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

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

---

<sup>5</sup> In addition to the recommended passing score *averaged* across the two panels, the recommended passing scores for *each* panel are presented.

---

# OVERVIEW OF THE *PRAXIS*<sup>®</sup> CORE ACADEMIC SKILLS FOR EDUCATORS: MATHEMATICS SUBTEST

The *Praxis*<sup>®</sup> Core Academic Skills for Educators: Mathematics (5733) *Study Companion* document (ETS, in press) describes the purpose and structure of the subtest. In brief, the *Praxis* Core Academic Skills for Educators subtests measure whether candidates entering a teacher preparation program have the necessary reading, writing, and mathematical knowledge/skills. Each subtest — Reading, Writing, and Mathematics — is administered and scored separately<sup>6</sup>.

The one hour thirty-minutes assessment contains 56 selected-response and numeric-entry items<sup>7</sup> covering three content areas: *Number and Quantity* (approximately 20 items), *Data Interpretation and Representation, Statistics, and Probability* (approximately 18 items), and *Algebra and Geometry* (approximately 18 items)<sup>8</sup> The reporting scale for the *Praxis* Core Academic Skills for Educators: Mathematics subtest ranges from 100 to 200 scale-score points.

---

## PROCESSES AND METHODS

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

The standard-setting study began as a general session for both panels. The session opened with a welcome and introduction by each of the meeting facilitators. The facilitators described the test, provided an overview of standard setting, and presented the agenda for the study. Appendix B shows the standard-setting study agenda.

### REVIEWING THE SUBTEST

While both panels were together during the general session, the standard-setting panelists took the test and then discussed the content measured. This discussion helped bring the panelists to a shared

---

<sup>6</sup> More details about the Reading (5712) and Writing (5722) subtests can be found on the [ETS website](#).

<sup>7</sup> Six of the 56 selected-response and numerical-entry items are pretest items and do not contribute to a candidate's score.

<sup>8</sup> The number of items for each content area may vary slightly from form to form of the test.

understanding of what the test does and does not cover, which serves to reduce potential judgment errors later in the standard-setting process.

The test discussion covered the major content areas being addressed by the subtest. Panelists were asked to remark on any content areas that would be particularly challenging for entry-level teachers or areas that address content particularly important for entry-level teachers.

## DEFINING THE JUST QUALIFIED CANDIDATE

Following the review of the subtest, 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 subtest score that aligns with this description.

Both panels worked together to create a description of the just qualified candidate — the knowledge/skills that differentiate a *just* from a *not quite* qualified candidate. To create this description, they first split into smaller groups to consider the just qualified candidate. Then they reconvened and, through whole-group discussion, created the description of the just qualified candidate to use for the remainder of the study.

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

## PANELISTS' JUDGMENTS

The standard-setting process for the *Praxis* Core Academic Skills for Educators: Mathematics subtest 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.

For consistency in understanding the standard-setting judgment process, both panels remained together as they received training and practice in how to complete their standard-setting judgments. 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. After the independent judgments were completed, panelists were split into two, distinct panels that worked separately for the remainder of the study.

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

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

---

## RESULTS

### EXPERT PANELS

Table 2 presents a summary of the panelists' demographic information across panels. The panel included 33 educators representing 20 states and Washington, DC. (See Appendix A for a listing of panelists.) 28 panelists were Faculty members responsible for instructing teacher preparation courses, three were administrators, and two held other positions. The number of experts by panel and their demographic information are presented in Appendix D (Table D1).



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

	<i>N</i>	<i>%</i>
<b>Current position</b>		
Educator Preparation Faculty	28	85
Program Administrator	3	9
Program Coordinator	1	3
Curriculum and Design Specialist	1	3
<b>Race/Ethnicity</b>		
White or European American	18	55
Black or African American	10	30
Hispanic or Latino	1	3
Asian or Asian American	1	3
Other	2	6
Prefer not to answer	1	3
<b>Gender</b>		
Female	25	76
Male	8	24
<b>Do you typically instruct courses that cover any of the following topics?</b>		
Assessment	22	67
Classroom Organization & Management	21	64
Comparative Education	5	15
Diversity	21	64
Education Technology	12	36
Families & Communities	11	33
Human Development & Learning	11	33
Introduction to Education	13	39
Inclusion	12	36
<b>Including this year, how many years of experience do you have preparing teacher candidates?</b>		
3 years or less	3	9
4 - 7 years	6	18
8 - 11 years	7	21
12 - 15 years	7	21
16 years or more	10	30

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

	<i>N</i>	<i>%</i>
<b>Do you typically instruct courses intended for teacher candidates across subject areas (e.g., mathematics, social studies) and grade levels?</b>		
Yes	28	85
No, I only instruct courses focused on a particular subject area but across grade levels (e.g., English/language arts pedagogy across grades K-12)	4	12
No, I only instruct courses focused on a particular subject area at a particular grade level (e.g., secondary science)	1	3
<b>Are you currently supervising or mentoring student teachers?</b>		
Yes	20	61
No	13	39

**STANDARD-SETTING JUDGMENTS**

Table 3 summarizes the standard-setting judgments (Round 2) of panelists. The table also includes estimates of the measurement error associated with the judgments: the standard deviation of the mean and the standard error of judgment (SEJ). The SEJ is one way of estimating the reliability or consistency of a panel’s standard-setting judgments.<sup>9</sup> It indicates how likely it would be for several other panels of educators similar in makeup, experience, and standard-setting training to the current panel to recommend the same passing score on the same form of the test. The confidence intervals created by adding/subtracting two SEJs to each panel’s recommended passing score overlap, indicating that they may be comparable.

Panelist-level results, for Rounds 1 and 2, are presented in Appendix D (Table D2).

<sup>9</sup> An SEJ assumes that panelists are randomly selected and that standard-setting judgments are independent. It is seldom the case that panelists are randomly sampled, and only the first round of judgments may be considered independent. The SEJ, therefore, likely underestimates the uncertainty of passing scores (Tannenbaum & Katz, 2013).

**Table 3**  
***Summary of Round 2 Standard-setting Judgments***

	<b>Panel 1</b>	<b>Panel 2</b>
Average	25.76	28.33
Lowest	17.70	23.60
Highest	34.65	34.70
SD	4.24	2.94
SEJ	1.03	0.74

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* Core Academic Skills for Educators: Mathematics test are 25.76 for Panel 1 and 28.33 for Panel 2 (out of a possible 50 raw-score points). The values were rounded to the next highest whole number, to determine the functional recommended passing score — 26 for Panel 1 and 29 for Panel 2. The scale scores associated with 26 and 29 raw points are 144 and 152, respectively.

In addition to the recommended passing score for each panel, the average passing score across the two panels is provided to help education agencies determine an appropriate passing score. The panels’ average passing score recommendation for the *Praxis* Core Academic Skills for Educators: Mathematics test is 27.05 (out of a possible 50 raw-score points). The value was rounded to 28 (next highest raw score) to determine the functional recommended passing score. The scale score associated with 28 raw points is 150.

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 Score<sup>10</sup>***

<b>Recommended passing score (CSEM)</b>		<b>Scale score equivalent</b>
	28 (3.55)	150
-2 CSEM	21	130
-1 CSEM	25	142
+ 1 CSEM	32	160
+ 2 CSEM	36	170

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

All panelists reported that the description of the just qualified candidate was at least *somewhat influential* in guiding their standard-setting judgments; 30 of the 33 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. Twenty-four of the 33 panelists indicated that their own professional experience was *very influential* in guiding their judgments.

All but one of the panelists indicated they were at least *somewhat comfortable* with the passing score they recommended; 29 of the 33 panelists were *very comfortable*. Thirty-one of the 33 panelists indicated the recommended passing score was *about right*, one of the remaining panelist indicated that the

<sup>10</sup> 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.

passing score was *too low* and the other panelist indicated the passing score was *too high*. The panelist who indicated that he or she was *very uncomfortable* with the recommended passing score indicated that it was *about right*.

---

## SUMMARY

To support the decision-making process for education agencies establishing a passing score (cut score) for the *Praxis* Core Academic Skills for Educators: Mathematics 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* Core Academic Skills for Educators: Mathematics test, the recommended passing score<sup>11</sup> is 28 out of a possible 50 raw-score points. The scale score associated with a raw score of 28 is 150 on a 100–200 scale.

---

<sup>11</sup> 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: Core Academic Skills for Educators: Mathematics (5733)*. 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***

---

<b><u>Panelist</u></b>	<b><u>Affiliation</u></b>
Ginger Behnke	University of Valley Forge (PA)
Joy Brown	University of North Alabama (AL)
Gunhan Caglayan	New Jersey City University (NJ)
Qiang Cheng	University of Mississippi (MS)
Nicholas Cluster	South Carolina State University (SC)
C. Neelie Dobbins	Southern Arkansas University (AR)
Sherita Flake	Johns Hopkins University (MD)
Keri Flowers	Troy University (AL)
Howard Gordon	University of Nevada, Las Vegas (NV)
LaChan Hannon	The College of New Jersey (NJ)
Tracy Hargrove	University of North Carolina Wilmington (NC)
Jean Hearn	Creighton University (NE)
Molly Hill	University of Louisiana Monroe (LA)
Sumitra Himangshu-Pennybacker	Middle Georgia State University (GA)
Sarah Hunt-Barron	University of South Carolina Upstate (SC)
Allen Jantz	Bethel College (KS)
Whitney Johnson	Morgan State University (MD)
Deena Khalil	Howard University (DC)
Gladys Labas	Southern CT State University (CT)
Paula Lucas	Marshall University (WV)
Kathy McDilda	West Virginia State University (WV)
Rebecca McMullen	Fort Valley State University (GA)

*(table continues on the next page)*



*Participating Panelists With Affiliation (continued)*

---

<b><u>Panelist</u></b>	<b><u>Affiliation</u></b>
Gloria Niles	University of Hawaii West Oahu (HI)
Chinenye Ofodile	Albany State University (GA)
Teresa Powell	Lincoln University (PA)
Catherine Schwartz	East Carolina University (NC)
Sandra Smith	Cumberland University (TN)
Tatia Totorica	Boise State University (ID)
Amy Vinlove	University of Alaska Fairbanks (AK)
Lisa Warner	William Paterson University (NJ)
Dennis Williams	Jackson State University (MS)
Meredith Wright	Henderson State University (AR)
Jamaal Young	University of Iowa (IA)

---

---

APPENDIX B  
STUDY AGENDA

# AGENDA

## ***Praxis*<sup>®</sup> Core Academic Skills for Educators: Mathematics (5733) Standard-Setting Study**

Day 1

Welcome and Introduction

Overview of Standard Setting and the *Praxis* Core Academic Skills for Educators: Mathematics Test

Review the *Praxis* Core Academic Skills for Educators: Mathematics Test

Break

Discuss the *Praxis* Core Academic Skills for Educators: Mathematics Test

Define the Knowledge/Skills of a Just Qualified Candidate

Lunch

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

Break

Training and practice in Modified Angoff Standard-setting judgments

Collect Materials; End of Day 1

# AGENDA

## ***Praxis*<sup>®</sup> Core Academic Skills for Educators: Mathematics (5733) Standard-Setting Study**

Day 2

Overview of Day 2

Review Training and introduce Tablets

Round 1 Standard-setting Judgments

Break

Round 1 Feedback & Round 2 judgments

Lunch

Feedback on Round 2 Recommended Passing Score

Complete Final Evaluation

Collect Materials; End of Study

---

## APPENDIX C

### JUST QUALIFIED CANDIDATE DESCRIPTION

## Description of the Just Qualified Candidate<sup>12</sup>

**A just qualified candidate ...**

### **Numbers & Quantity**

1. Solve problems involving integers, decimals, percents and fractions.
2. Demonstrate an understanding of place value, naming of decimal numbers, and ordering of numbers.
3. Can translate basic word problems into numerical problems and/or models.
4. Solve contextual problems by identifying relevant numbers, information, or operations including identification of counter examples to statements using basic arithmetic and applying order of operations.

### **Data Interpretation and Representation, Statistics, and Probability**

5. Solve problems involving basic statistics including mean, median, mode, and range.
6. Compute simple probabilities and use probabilities to solve simple problems.
7. Identify and interpret positive and negative linear relationships when represented graphically or algebraically.

### **Algebra**

8. Write a two-step equation or expression that models a real-life or mathematical problem.
9. Solve and/or model contextual problems that can include linear relationships.
10. Solve single variable, multi-step linear equations with context and/or without context.
11. Use properties of operations to simplify and evaluate algebraic expressions.

### **Geometry**

12. Use properties of angles and common two dimensional shapes to solve problems.

---

<sup>12</sup> 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)***

	<b>Panel 1</b>		<b>Panel 2</b>	
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
<b>Current position</b>				
Educator Preparation Faculty	14	82	14	88
Program Administrator	2	12	1	6
Program Coordinator	1	6	0	0
Curriculum and Design Specialist	0	0	1	6
<b>Race</b>				
White or European American	10	59	8	50
Black or African American	5	29	5	31
Hispanic or Latino	1	6	0	0
Asian or Asian American	0	0	1	6
Other	0	0	2	13
Prefer not to answer	1	6	0	0
<b>Gender</b>				
Female	13	76	12	75
Male	4	24	4	25
<b>Do you typically instruct courses that cover any of the following topics?</b>				
Assessment	10	59	12	75
Classroom Organization & Management	11	65	10	63
Comparative Education	3	18	2	13
Diversity	10	59	11	69
Education Technology	7	41	5	31
Families & Communities	5	29	6	38
Human Development & Learning	6	35	5	31
Introduction to Education	7	41	6	38
Inclusion	4	24	8	50
<b>Including this year, how many years of experience do you have preparing teacher candidates?</b>				
3 years or less	1	6	2	13
4 - 7 years	2	12	4	25
8 - 11 years	3	18	4	25
12 - 15 years	4	24	3	19
16 years or more	7	41	3	19



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

	<b>Panel 1</b>		<b>Panel 2</b>	
	<i>N</i>	%	<i>N</i>	%
<b>Do you typically instruct courses intended for teacher candidates across subject areas (e.g., mathematics, social studies) and grade levels?</b>				
Yes	13	76	15	94
No, I only instruct courses focused on a particular subject area but across grade levels (e.g., English/language arts pedagogy across grades K-12)	3	18	1	6
No, I only instruct courses focused on a particular subject area at a particular grade level (e.g., secondary science)	1	6	0	0
<b>Are you currently supervising or mentoring student teachers</b>				
Yes	11	65	9	56
No	6	35	7	44

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

---

<b>Panelist</b>	<b>Panel 1</b>		<b>Panel 2</b>	
	<b>Round 1</b>	<b>Round 2</b>	<b>Round 1</b>	<b>Round 2</b>
1	27.50	27.50	30.75	30.65
2	22.65	24.05	24.50	23.60
3	28.90	28.50	28.35	28.75
4	34.70	30.75	30.10	29.50
5	23.70	24.50	31.60	31.20
6	32.10	29.70	30.50	28.10
7	29.65	25.30	33.10	33.10
8	25.70	26.45	23.25	25.85
9	15.00	17.70	27.70	27.00
10	24.20	22.10	24.00	26.05
11	33.55	34.65	35.80	34.70
12	18.50	20.20	26.80	25.00
13	23.70	23.40	27.80	27.50
14	26.70	25.10	26.30	26.70
15	32.20	28.60	28.60	28.40
16	21.20	21.10	27.25	27.20
17	28.90	28.40		
<b>Average</b>	26.40	25.76	28.53	28.33
<b>Lowest</b>	15.00	17.70	23.25	23.60
<b>Highest</b>	34.70	34.65	35.80	34.70
<b>SD</b>	5.38	4.24	3.37	2.94
<b>SEJ</b>	1.31	1.03	0.84	0.74

---

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

	<b>Strongly agree</b>		<b>Agree</b>		<b>Disagree</b>		<b>Strongly disagree</b>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
• I understood the purpose of this study.	17	100	0	0	0	0	0	0
• The instructions and explanations provided by the facilitators were clear.	17	100	0	0	0	0	0	0
• The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	17	100	0	0	0	0	0	0
• The explanation of how the recommended passing score is computed was clear.	15	88	2	12	0	0	0	0
• The opportunity for feedback and discussion between rounds was helpful.	17	100	0	0	0	0	0	0
• The process of making the standard-setting judgments was easy to follow.	15	88	2	12	0	0	0	0
• I understood how to use the survey software.	17	100	0	0	0	0	0	0

**Table D3 (continued)**  
**Final Evaluation: Panel 1**

<b>How influential was each of the following factors in guiding your standard-setting judgments?</b>	<b>Very influential</b>		<b>Somewhat influential</b>		<b>Not influential</b>			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• The description of the just qualified candidate	15	88	2	12	0	0		
• The between-round discussions	14	82	3	18	0	0		
• The knowledge/skills required to answer each test item	15	88	2	12	0	0		
• The passing scores of other panel members	6	35	11	65	0	0		
• My own professional experience	15	88	2	12	0	0		
	<b>Very comfortable</b>		<b>Somewhat comfortable</b>		<b>Somewhat uncomfortable</b>		<b>Very uncomfortable</b>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
• Overall, how comfortable are you with the panel's recommended passing score?	15	88	2	12	0	0	0	0
	<b>Too low</b>		<b>About right</b>		<b>Too high</b>			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• Overall, the recommended passing score is:	1	6	15	88	1	6		

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

	<b>Strongly agree</b>		<b>Agree</b>		<b>Disagree</b>		<b>Strongly disagree</b>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
• I understood the purpose of this study.	15	94	1	6	0	0	0	0
• The instructions and explanations provided by the facilitators were clear.	10	63	6	38	0	0	0	0
• The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	12	75	4	25	0	0	0	0
• The explanation of how the recommended passing score is computed was clear.	15	94	1	6	0	0	0	0
• The opportunity for feedback and discussion between rounds was helpful.	12	75	4	25	0	0	0	0
• The process of making the standard-setting judgments was easy to follow.	12	75	4	25	0	0	0	0
• I understood how to use the survey software.	16	100	0	0	0	0	0	0

**Table D4 (continued)**  
**Final Evaluation: Panel 2**

<b>How influential was each of the following factors in guiding your standard-setting judgments?</b>	<b>Very influential</b>		<b>Somewhat influential</b>		<b>Not influential</b>			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• The description of the just qualified candidate	15	94	1	6	0	0		
• The between-round discussions	10	63	6	38	0	0		
• The knowledge/skills required to answer each test item	10	63	6	38	0	0		
• The passing scores of other panel members	5	31	9	56	2	13		
• My own professional experience	9	56	7	44	0	0		
	<b>Very comfortable</b>		<b>Somewhat comfortable</b>		<b>Somewhat uncomfortable</b>		<b>Very uncomfortable</b>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
• Overall, how comfortable are you with the panel's recommended passing score?	14	88	1	6	0	0	1	6
	<b>Too low</b>		<b>About right</b>		<b>Too high</b>			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• Overall, the recommended passing score is:	0	0	16	100	0	0		