Multistate Standard-Setting Technical Report for the *Praxis*® Special Education: Foundational Knowledge (5355)

Student and Teacher Assessments: Validity and Test Use

ETS

Princeton, New Jersey

January 2023

Executive Summary

To support the decision-making process of education agencies establishing a passing score (cut score) for the *Praxis*® Special Education: Foundational Knowledge (5355) test, research staff from Educational Testing Service (ETS) designed and conducted a multistate standard-setting study (Tannenbaum, 2011, 2012).

Participating States

Panelists from 14 states were recommended by their respective education agencies. The education agencies recommended panelists with (a) experience as special education teachers or college faculty who prepare those special education teachers and (b) familiarity with the knowledge and skills required of beginning special education teachers.

Recommended Passing Score

ETS provides a recommended passing score from the multistate standard-setting study to help education agencies determine an appropriate operational passing score. For the *Praxis* Special Education: Foundational Knowledge test, the recommended passing score[1] is 62 out of a possible 110 raw-score points. The scale score associated with a raw score of 62 is 145 on a 100–200 scale.

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

Introduction

To support the decision-making process for education agencies establishing a passing score (cut score) for the *Praxis*® Special Education: Foundational Knowledge (5355) test, research staff from ETS designed and conducted a multistate standard-setting study (Tannenbaum, 2011, 2012) in December 2022. Education agencies[2] recommended panelists with (a) experience as either special education teachers or (b) familiarity with the knowledge and skills required of beginning special education teachers. Fourteen states (Table 1) were represented by 29 panelists. (See Appendix A for the names and affiliations of the panelists.)

Table 1
Participating States and the Number of Panelists

Arkansas (4 panelists)	Mississippi (2 panelists)
Colorado (2 panelists)	Nebraska (1 panelist)
Delaware (3 panelists)	Nevada (1 panelist)
Idaho (1 panelist)	Rhode Island (1 panelist)
Kansas (2 panelists)	Tennessee (3 panelists)
Kentucky (3 panelists)	West Virginia (1 panelist)
Maryland (1 panelist)	Wyoming (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, 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, which represents the combined judgments of a group of experienced educators. Each state may want to consider the recommended passing score but also other sources of information when setting the final *Praxis* Special Education: Foundational Knowledge passing score (see Geisinger & McCormick, 2010). A state may

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

accept the recommended passing score, adjust the score upward to reflect more stringent expectations, or adjust the score downward to reflect more lenient expectations. There is no *correct* decision; the appropriateness of any adjustment may only be evaluated in terms of its meeting the state's needs.

Two sources of information to consider when setting the passing score are the standard error of measurement (SEM) and the standard error of judgment (SEJ). The former addresses the reliability of the *Praxis* Special Education: Foundational Knowledge test score and the latter, the reliability of panelists' passing-score recommendation. The SEM allows states to recognize that any test score on any standardized test—including a *Praxis* Special Education: Foundational Knowledge 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 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 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 they should receive a license/certificate, but their 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 they should not receive a license/certificate, but they actually do possess the required knowledge/skills. States needs to consider which decision error is more important to minimize.

Overview of the *Praxis*® Special Education: Foundational Knowledge Test

The *Praxis*® Special Education: Foundational Knowledge *Study Companion* document (ETS, in press) describes the purpose and structure of the test. In brief, the test measures whether entry-level

special education teachers have the knowledge/skills believed necessary for competent professional practice.

The two-hour assessment contains 120 selected-response items[3] covering four content areas: *Human Development and Individual Learning Differences* (approximately 32 items), *Effective Planning and Instruction and Productive Learning Environments* (approximately 38 items), *Assessment* (approximately 27 items), and *Professional Learning, Practice, and Collaboration* (approximately 23 items).[4] The reporting scale for the *Praxis* Special Education: Foundational Knowledge test 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.

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

Reviewing the Test

The standard-setting panelists first took the test and then discussed the content measured. This discussion helped bring the panelists to a shared understanding of what the test does and does not cover, which serves to reduce potential judgment errors later in the standard-setting process.

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 special education teachers or areas that address content particularly important for entry-level special education teachers.

[3] Ten of the 120 selected-response items are pretest items and do not contribute to a candidate's score.
[4] The number of items for each content area may vary slightly from form to form of the test.

Defining the Just-Qualified Candidate

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

The panelists created a description of the just-qualified candidate, focusing on the knowledge/skills that differentiate a *just-qualified* from a *not quite-qualified* candidate. To create this description, the panelists from both panels were assigned to three smaller groups in order to create a draft description. Then they reconvened and, through whole-group discussion of the three drafts, reached consensus on to determine the final version. This final description of the just-qualified candidate was used by both panels for the remainder of the study.

The description of the just-qualified candidate summarized the panels' discussion in a list 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. A clean, PDF-version of the final description was distributed to panelists to use for the remaining phases of the study (see Appendix C for the just-qualified candidate description).

Given that the two-panel multistate standard-setting study was designed to provide two recommendations for the same performance standard, it was important that panels use a consistent just-qualified candidate description to frame their judgments. Therefore, the panelists from both panels worked together until the just-qualified candidate description was finalized.

Panelists' Judgments

The standard-setting process for the *Praxis* Special Education: Foundational Knowledge test was a probability-based Modified Angoff method (Brandon, 2004; Hambleton & Pitoniak, 2006). Using this method, 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 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 in the Modified Angoff method 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. Item-level data were highlighted to show when panelists converged in their judgments or diverged in their judgments (i.e., when at least two-thirds of the panelists' judgments were in the same difficulty range).

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

In Round 2, panelists discussed their Round 1 judgments and were encouraged by the facilitator (a) to share the rationales for their judgments and (b) to consider their judgments in light of the rationales provided by the other panelists. Panelists recorded their Round 2 judgments only for items

when they wished to change a Round 1 judgment. Panelists' final judgments for the study, therefore, consist of their Round 1 judgments and any adjusted judgments made during Round 2.

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

Results

Expert Panels

Table 2 presents a summary of the panelists' demographic information. The panel included 29 educators representing 14 states. (See Appendix A for a listing of panelists.) Eleven panelists were special education teachers and two were educational specialists. Of the three administrators and department heads, one was a district-level educational leader and two were school department heads. Eleven panelists were college faculty and two held other positions. Ten of the eleven of the faculty members' job responsibilities included the training of special education teachers. The number of experts by panel and their demographic information are presented in Appendix D (Table D1).

Table 2
Panel Member Demographics (Across Panels)

Background Survey Question	Number	Percent
What is your current position?	**N**	**%**
Special education teacher	11	38
Educational specialist	2	7
Administrator or Department Head	3	10
College faculty	11	38
Other: Special education teacher (9-12th) and department head/LSC	1	3
Other: Retired director of special education for Gallatin County Schools	1	3
How do you describe yourself (i.e., race/ethnicity)?	**N**	**%**
Asian or Asian American	1	3
Black or African American	9	31
White	17	59
Other: Hispanic or Latino/Middle Eastern or North African/White	1	3
Prefer not to respond	1	3

Table 2 (continued from the previous page)
Panel Member Demographics (Across Panels)

What is your gender?	**N**	**%**
Female/Woman	23	79
Male/Man	5	17
Non-binary	0	0
Prefer to self-describe	0	0
Prefer not to respond	1	3
Are you currently certified as a special education teacher in your state?	**N**	**%**
Yes	18	62
No	0	0
I am not currently working at the P-12 level	11	38
Are you currently teaching special education in your state?	**N**	**%**
Yes	12	41
No	6	21
I am not currently working at the P-12 level	11	38
Are you currently supervising or mentoring other special education teachers?	**N**	**%**
Yes	15	52
No	3	10
I am not currently working at the P-12 level	11	38
At what P–12 grade level are you currently teaching special education?	**N**	**%**
Elementary (P - 5 or P - 6)	2	7
Elementary and middle school	1	3
Middle school (6 - 8 or 7 - 9)	4	14
Middle and high school	1	3
High school (9–12 or 10–12)	3	10
Other	3	10
Not currently teaching at the P–12 level	15	52
Including this year, how many years of experience do you have teaching special education?	**N**	**%**
3 years or less	1	3
4–7 years	10	34
8–11 years	2	7
12–15 years	1	3
16 years or more	4	14
I am not currently working at the P-12 level	11	38
Which best describes the location of your P–12 school?	**N**	**%**
Urban	5	17
Suburban	4	14
Rural	6	21
Not working in a school (e.g., district-level)	3	10

Table 2 (continued from the previous page)
Panel Member Demographics (Across Panels)

If you are college faculty, are you currently involved in the training/ preparation of special education teachers?	**N**	**%**
Yes	10	34
No	1	3
Not college faculty	18	62

Standard-Setting Judgments

Panelist-level results, for Rounds 1 and 2, are presented in Appendix D (Tables D2 – D4). The mean represents the panel's passing score recommendation after Round 2. Table 3 also includes the standard deviation 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. 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. (Appendix E provides the technical notes, which further describe the SEJ.)

Table 3
Summary of Round 2 Standard-setting Judgments by Panel

Statistic	**Panel 1**	**Panel 2**
Mean	60.61	63.22
Minimum	45.90	53.35
Maximum	68.55	76.95
SD	5.82	6.18
SEJ	1.50	1.65

Data from Panelist 11 on panel 1 and Panelist 2 on panel 2 were detected to be outliers (High, 2000; see Appendix E). However, ETS does not recommend that their data be removed from the panel recommendation. Based on a report from the panel facilitators, the panelists were believed to be following the standard-setting process faithfully. Throughout the standard-setting, panelists are encouraged to consider the perspectives of their colleagues but that were not required to agree with their judgments.

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 4).

The Round 2 mean score is the panel's final recommended passing score. The panel's passing score recommendation for the *Praxis* Special Education: Foundational Knowledge test are 60.61 for Panel 1 and 63.22 for Panel 2 (out of a possible 110 raw-score points). The values were rounded to the next highest whole number to determine the functional recommended passing score--61 for Panel 1 and 64 for Panel 2. The scale scores associated with 61 and 64 raw points are 144 and 148, 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* Special Education: Foundational Knowledge test is 61.92 (out of a possible 110 raw-score points). The value was rounded to 62 (next highest raw score) to determine the functional recommended passing score. The scale score associated with 62raw points is 145.

The conditional standard error of measurement (CSEM) around the recommended passing score is 5.23 raw points. A standard error represents the uncertainty associated with a test score (See Appendix E for further information about the CSEM.) Table 5 shows the raw scores and the scale scores associated with one and two CSEM below and above the recommended passing score.

Table 5

Scores 1 and 2 CSEM Around the Recommended Passing Score (RPS)

Scores	Raw Score Points out of 110	Praxis Scale Score Equivalent
RPS - 2 CSEM	52	132
RPS - 1 CSEM	57	139
RPS	**62**	**145**
RPS +1 CSEM	68	153
RPS +2 CSEM	73	159

Notes. CSEM = conditional standard error(s) of measurement. The CSEM of the recommended passing score is 5.23 raw points. 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 then converted to scale scores.

Final Evaluations

The panelists completed an evaluation at the conclusion of the 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 shown the panel's recommended passing score after Round 2 and asked, in the evaluation, (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, per panel, is presented in Appendix D (Tables D5 – D14).

All panelists *strongly agreed* that they understood the purpose of the study. All panelists *strongly agreed* or *agreed* that the facilitator's instructions and explanations were clear. All panelists *strongly agreed* or *agreed* that they were prepared to make their standard-setting judgments. 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. All of the panelists reported that between-round discussions were at least *somewhat influential* in guiding their judgments. Eighteen of the 29 panelists) indicated that their own professional experience was *very influential* in guiding their judgments.

Twenty-four of the 29 panelists indicated they were *very comfortable* with the passing score they recommended; the remaining panelists indicated they were *somewhat comfortable* with the recommended passing score. Twenty-seven of the 29 panelists indicated the recommended passing score was *about right*; the remaining two panelists felt that the panel's recommended score was *too low*.

Summary

To support the decision-making process for education agencies establishing a passing score (cut score) for the *Praxis* Special Education: Foundational Knowledge 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* Special Education: Foundational Knowledge test, the recommended passing score[5] is 62 out of a possible 110 raw-score points. The scale score associated with a raw score of 62 is 145 on a 100–200 scale.

[5] 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.

Brennan, R. L. (2002, October). Estimated standard error of a mean when there are only two observations (Center for Advanced Studies in Measurement and Assessment Technical Note Number 1). *Iowa City: University of Iowa*.

Cizek, G. J., & Bunch, M.B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage.

ETS. (in press). *The Praxis Series®: The Praxis Study Companion: Special Education: Foundational Knowledge (5355)*. 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.

Lord. F.M. (1984). Standard errors of measurement at different ability levels. *Journal of Educational Measurement, 21*, 239-283.

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.

Tannenbaum, R. J. (2011). Setting standards on *The Praxis Series*™ tests: A multistate approach. *R&D Connections*, *17*, 1-9.

Tannenbaum, R. J. (2012). A multistate approach to setting standards: An application to teacher licensure tests. *CLEAR Exam Review*, *23*(1), 18-24.

Appendix A: Panelists' Names & Affiliations

Participating Panelists With Affiliation and State

Panelist Name	**Panelists' Affiliation and State Abbreviation**
Bret Arwood	Sweetwater County School District #1 (WY)
Valerie Bahige	Campbell County School District (WY)
Shannon Banashak	Biloxi Public School District (MS)
Caitlin Best	Colonial School District (DE)
Michelle Bills	Eastside Elementary (WY)
Jerrie Brooks	Fort Hays State University (KS)
Tammaniqua Carr	Canton Public School District (MS)
Hayden Cook	Bethany College (WV)
Carol Douglass	Harding University (AR)
Veronica Fiedler	Colorado Department of Education (CO)
Terrance Hannah	Hamilton County School System (TN)
Frederick Hoppe	Lincoln Public Schools (RI)
Jennifer Hune	UA Little Rock (AR)
Jessica Jersild	Morrill Public Schools (NE)
Jo Ann McCaughan	Retired Director of Special Education from Gallatin County Schools (KY)
Aundrea McFall	Freed-Hardeman University (TN)
Makenzie Meade	West Grand School District (CO)
Ma Naneth Mendoza	Reed Elementary School (AR)
Barbara Mitchell	Kansas State University (KS)
Kimberly Noah	Great Basin College (NV)
Shannon Patrick	Baltimore city Public Schools (MD)
Jacee Phillips	UW Lab School (WY)
Kimberly Puckett	Model Lab School at Eastern Kentucky University (KY)
Kristie Robinette	Sullivan County Schools (TN)

Panelist Name	Panelists' Affiliation and State Abbreviation
Megan Schnetzer	Wilmington University (DE)
Kayla Steltenkamp	Thomas More University (KY)
Jahsha Tabron	Delaware Professional Standards Board (DE)
Royal Toy	Lewis-Clark State College (ID)
Vashanti Williams	Hot Springs School District (AR)

Appendix B: Agenda

Praxis® Special Education: Foundational Knowledge (5355)

Standard-Setting Study

MONDAY, DECEMBER 5, 2022

10:00 AM EST/9:00 AM CST/ 8:00 AM MST/7:00 AM PST	Welcome, introductions, and goals for the day Standard-setting overview presentation • Q&A about the training • Polling: Checking for understanding Test familiarization for the test • Overview & instructions • Independent test familiarization • Self-scoring instructions • Independent self-scoring **Break** Discussion of the content measured **Lunch break** Just-Qualified Candidate (JQC) • Overview • Polling: Who is the JQC? • Drafting the JQC description in small groups **Break as needed in small groups** Preparation for Day 2
7:00 PM EST/ 6:00 PM CST/ 5:00 PM MST/ 4:00 PM PST	End of Day 1

Praxis® Special Education: Foundational Knowledge (5355)

Standard-Setting Study

TUESDAY, DECEMBER 6, 2022

10:00 AM EST/9:00 AM CST/ 8:00 AM MST/7:00 AM PST	Overview of the Day Just-Qualified Candidate (JQC) (continued) • Finalizing the JQC description – whole group consensus **Break** Resume in Separate panels Standard Setting Training for Selected-Response Items • Instructions and materials • Independent practice round judgments **LUNCH BREAK** Practice Round Data Discussion • Instructions, materials, and screen setup • Discussion of the practice round data • Polling: Evaluation of the judgment training Round 1 Standard Setting Judgments **Break individually as needed** Check out before completing work for the day • Save and send the judgment form (instructions provided) • Facilitator will confirm the data is received by each person
7:00 PM EST/ 6:00 PM CST/ 5:00 PM MST/ 4:00 PM PST	End of Day 2

Praxis® Special Education: Foundational Knowledge (5355)

Standard-Setting Study

WEDNESDAY, DECEMBER 7, 2022

10:00 AM EST/9:00 AM CST/ 8:00 AM MST/7:00 AM PST	Overview of the Day
	
	Honoraria Payment Process Presentation
	
	Round 1 feedback: Summary data
	• Polling: Evaluation of data presentation
	
	Round 1 feedback: Item-level data and Round 2 judgments (**break as needed**)
	• Check out when finished
	
	LUNCH BREAK
	
	Round 2 feedback: Recommended Passing Scores
	
	Complete final evaluation
	
	Wrap Up/ Final Steps
	• Review Nondisclosure of Secure Materials
	  ○ Destruction of Files
	  ○ What you can/cannot discuss
	• Final Questions and Thank You

4:00 PM EST/ 3:00 PM CST/ 2:00 PM MST/ 1:00 PM PST	End of Study

Appendix C: Just-Qualified Candidate Description

Description of the Just-Qualified Candidate[6]

A just-qualified candidate…

1. Understands typical and atypical human development and behavior and how these can impact student learning
2. Understands the basic characteristics and defining factors for each of the major disability categories under IDEA
3. Knows the basic characteristics of designing coherent instruction and how to adapt them to diverse learners (e.g., students with disabilities, English Language learners, etc.) and environments
4. Is familiar with how to gather and use data to design, implement, and adjust positive behavior interventions and support systems for individuals, groups, and classroom management
5. Is familiar with and attempts to implement research-based approaches to differentiating and delivering instruction to meet individual needs
6. Knows how to select and is familiar with administering and interpreting appropriate assessments to make informed data-driven decisions
7. Knows the basics of MTSS, including the role of RTI
8. Understands the role of the special education teacher in the pre-referral process.
9. Knows the major components of IDEA and other relevant legal mandates
10. Knows how to initiate appropriate communication and collaboration with families and school and community professionals
11. Is familiar with the roles of support professionals in relation to individual students (e.g., related service providers, paraprofessionals, general education teachers, administration)
12. Knows the required components of an IEP

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

Appendix D: Panel-Specific Results

Table D1

Panel Member Demographics per Panel

Background Survey Question	Panel 1 Number	Panel 1 Percent	Panel 2 Number	Panel 2 Percent
What is your current position?	**N**	**%**	**N**	**%**
Special education teacher	6	40	5	36
Educational specialist	2	13	0	0
Administrator or Department Head	1	7	2	14
College faculty	5	33	6	43
Other: Special education teacher (9-12th) and department head/LSC	1	7	0	0
Other: Retired director of special education for Gallatin County Schools	0	0	1	7
How do you describe yourself (i.e., race/ethnicity)?	**N**	**%**	**N**	**%**
Asian or Asian American	1	7	0	0
Black or African American	5	33	4	29
White	8	53	9	64
Other: Hispanic or Latino/Middle Eastern or North African/White	0	0	1	7
Prefer not to respond	1	7	0	0
What is your gender?	**N**	**%**	**N**	**%**
Female/Woman	11	73	12	86
Male/Man	3	20	2	14
Non-binary	0	0	0	0
Prefer to self-describe	0	0	0	0
Prefer not to respond	1	7	0	0
Are you currently certified as a special education teacher in your state?	**N**	**%**	**N**	**%**
Yes	10	67	8	57
No	0	0	0	0
I am not currently working at the P-12 level	5	33	6	43

Table D1 (continued from previous page)
Panel Member Demographics per Panel

Background Survey Question	Panel 1 Number	Panel 1 Percent	Panel 2 Number	Panel 2 Percent
Are you currently teaching special education in your state?	**N**	**%**	**N**	**%**
Yes	7	47	5	36
No	3	20	3	21
I am not currently working at the P-12 level	5	33	6	43
Are you currently supervising or mentoring other special education teachers?	**N**	**%**	**N**	**%**
Yes	9	60	6	43
No	1	7	2	14
I am not currently working at the P-12 level	5	33	6	43
At what P–12 grade level are you currently teaching special education?	**N**	**%**	**N**	**%**
Elementary (P - 5 or P - 6)	2	13	0	0
Elementary and middle school	1	7	0	0
Middle school (6 - 8 or 7 - 9)	1	7	3	21
Middle and high school	1	7	0	0
High school (9–12 or 10–12)	2	13	1	7
Other	1	7	2	14
Not currently teaching at the P–12 level	7	47	8	57
Including this year, how many years of experience do you have teaching special education?	**N**	**%**	**N**	**%**
3 years or less	0	0	1	7
4–7 years	7	47	3	21
8–11 years	1	7	1	7
12–15 years	0	0	1	7
16 years or more	2	13	2	14
I am not currently working at the P-12 level	5	33	6	43

Table D1 (continued from previous page)
Panel Member Demographics per Panel

Background Survey Question	Panel 1 Number	Panel 1 Percent	Panel 2 Number	Panel 2 Percent
Which best describes the location of your K–12 school?	**N**	**%**	**N**	**%**
Urban	3	20	2	14
Suburban	1	7	3	21
Rural	4	27	2	14
Not working in a school (e.g., district-level)	2	13	1	7
Not currently working at the P–12 level	5	33	6	43
If you are college faculty, are you currently involved in the training/ preparation of special education teachers?	**N**	**%**	**N**	**%**
Yes	5	33	5	36
No	0	0	1	7
Not college faculty	10	67	8	57

Table D2

Panel 1 Passing Score Summary by Round of Judgments

Panelist	Round 1	Round 2
1	60.60	58.10
2	62.50	61.40
3	74.20	68.55
4	63.95	63.45
5	53.00	54.80
6	59.60	58.90
7	57.05	57.05
8	56.40	59.05
9	59.00	61.30
10	56.20	59.40
11	45.70	45.90
12	68.40	66.70
13	67.05	65.80
14	60.80	60.30
15	71.70	68.50

Table D3

Panel 2 Passing Score Summary by Round of Judgments

Panelist	Round 1	Round 2
1	59.35	59.45
2	78.85	76.95
3	60.55	62.05
4	72.05	69.10
5	78.70	64.30
6	65.45	64.05
7	66.40	66.40
8	70.00	68.80
9	52.30	54.10
10	56.40	58.20
11	61.40	60.40
12	71.50	63.60
13	65.40	64.30
14	50.15	53.35

Table D4

Summary of Standard-setting Judgments by Panel and by Round

Statistic	Panel 1, Round 1	Panel 1, Round 2	Panel 2, Round 1	Panel 2, Round 2
Mean	61.08	60.61	64.89	63.22
Minimum	45.70	45.90	50.15	53.35
Maximum	74.20	68.55	78.85	76.95
SD	7.34	5.82	8.85	6.18
SEJ	1.90	1.50	2.37	1.65

Table D5: Panel 1 Final Evaluation Process Questions

Likert Statement	Strongly agree *N*	Strongly agree %	Agree *N*	Agree %	Disagree *N*	Disagree %	Strongly disagree *N*	Strongly disagree %
I understood the purpose of this study.	15	100	0	0	0	0	0	0
The instructions and explanations provided by the facilitators were clear.	15	100	0	0	0	0	0	0
The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	14	93	1	7	0	0	0	0
The explanation of how the recommended passing score is computed was clear.	9	60	6	40	0	0	0	0
The opportunity for feedback and discussion for round 2 judgments was helpful.	15	100	0	0	0	0	0	0
The process of making the standard-setting judgments was easy to follow.	14	93	1	7	0	0	0	0

Table D6: Panel 1 Final Evaluation: Standard-Setting Process

	Too much time *N*	Too much time %	About the right amount of time *N*	About the right amount of time %	Too little time *N*	Too little time %
Small group JQC drafts	1	7	14	93	0	0
Whole group JQC consensus	1	7	13	87	1	7
Training and practice for making standard-setting judgments	0	0	13	87	2	13
Round 1 judgments (independent)	0	0	13	87	2	13
Round 2 judgments (with discussion)	1	7	14	93	0	0

Table D7: Panel 1 Final Evaluation: Influences in Standard-Setting Judgments

How influential was each of the following factors in guiding your standard-setting judgments?	Very influential *N*	Very influential %	Somewhat influential *N*	Somewhat influential %	Not influential *N*	Not influential %
The description of the just-qualified candidate	13	87	2	13	0	0
The between-round discussions	9	60	6	40	0	0
The knowledge/skills required to answer each test item	11	73	4	27	0	0
The passing scores of other panel members	3	20	10	67	2	13
My own professional experience	10	67	5	33	0	0

Table D8: Panel 1 Final Evaluation: Comfort with the Panel's Recommendation

Question	Very comfort-able *N*	Very comfort-able %	Somewhat comfort-able *N*	Somewhat comfort-able %	Somewhat uncom-fortable *N*	Somewhat uncom-fortable %	Very uncom-fortable *N*	Very uncom-fortable %
Overall, how comfortable are you with the panel's recommended passing score?	12	80	3	20	0	0	0	0

Table D9: Panel 1 Final Evaluation: Opinion of the Final Recommendation

Statement	Too low *N*	Too low %	About right *N*	About right %	Too high *N*	Too high %
Overall, the recommended passing score is:	1	7	14	93	0	0

Table D10: Panel 2 Final Evaluation Process Questions

Likert Statement	Strongly agree *N*	Strongly agree %	Agree *N*	Agree %	Disagree *N*	Disagree %	Strongly disagree *N*	Strongly disagree %
I understood the purpose of this study.	14	100	0	0	0	0	0	0
The instructions and explanations provided by the facilitators were clear.	13	93	1	7	0	0	0	0
The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	14	100	0	0	0	0	0	0
The explanation of how the recommended passing score is computed was clear.	13	93	1	7	0	0	0	0
The opportunity for feedback and discussion for round 2 judgments was helpful.	14	100	0	0	0	0	0	0
The process of making the standard-setting judgments was easy to follow.	12	86	2	14	0	0	0	0

Table D11: Panel 2 Final Evaluation: Standard-Setting Process

	Too much time *N*	Too much time %	About the right amount of time *N*	About the right amount of time %	Too little time *N*	Too little time %
Small group JQC drafts	1	7	12	86	1	7
Whole group JQC consensus	0	0	13	93	1	7
Training and practice for making standard-setting judgments	0	0	14	100	0	0
Round 1 judgments (independent)	1	7	13	93	0	0
Round 2 judgments (with discussion)	0	0	14	100	0	0

Table D12: Panel 2 Final Evaluation: Influences in Standard-Setting Judgments

How influential was each of the following factors in guiding your standard-setting judgments?	Very influential *N*	Very influential %	Somewhat influential *N*	Somewhat influential %	Not influential *N*	Not influential %
The description of the just-qualified candidate	14	100	0	0	0	0
The between-round discussions	11	79	3	21	0	0
The knowledge/skills required to answer each test item	10	71	4	29	0	0
The passing scores of other panel members	3	21	10	71	1	7
My own professional experience	8	57	5	36	1	7

Table D13: Panel 2 Final Evaluation: Comfort with the Panel's Recommendation

Question	Very comfort-able *N*	Very comfort-able %	Somewhat comfort-able *N*	Somewhat comfort-able %	Somewhat uncom-fortable *N*	Somewhat uncom-fortable %	Very uncom-fortable *N*	Very uncom-fortable %
Overall, how comfortable are you with the panel's recommended passing score?	12	86	2	14	0	0	0	0

Table D14: Panel 2 Final Evaluation: Opinion of the Final Recommendation

Statement	Too low *N*	Too low %	About right *N*	About right %	Too high *N*	Too high %
Overall, the recommended passing score is:	1	7	13	93	0	0

Appendix E: Technical Notes

Standard Error of Judgment (SEJ)

The standard error of judgment (SEJ) is one way of estimating the reliability or consistency of a panel's standard-setting judgments. 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 threshold score on the same form of the assessment. The 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 threshold scores (Tannenbaum & Katz, 2013).

The SEJ is calculated by dividing the standard deviation of the panelists' judgments (*SD*) by the square root of the number of panelists (*n*). The result serves as an estimate of the standard error of the mean (Brennan, 2002).

$$SEJ = \mathrm{SD} / \sqrt{n}$$

Outlier Analysis

An analysis of the data is conducted per panel. Judgments that are above or below 1.5 times the interquartile range for that panel are identified as outliers (High, 2000). ETS makes recommendations on the removal of *specific* outliers based on the observations of the panel facilitator. The panel facilitator reports whether or not the specified panelist was faithfully participating in the standard-setting process. The decision to accept the panel recommendation with or without the outlier data is solely at the discretion of the state.

Conditional Standard Error of Measurement (CSEM)

The conditional standard error of measurement (*CSEM*) for a test is computed from the study value (*SV*) of the recommended passing score and the number of selected-response items (*n*) on the test (see Lord, 1984):

$$CSEM = \sqrt{(SV)(n - SV)/(n - 1)}$$