



Multistate Standard-Setting Technical Report

ETS[®] SCHOOL SUPERINTENDENT ASSESSMENT (6991)

Educational Testing Service

Princeton, New Jersey

January 2019

EXECUTIVE SUMMARY

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

PARTICIPATING STATES

Panelists from 11 states were recommended by their respective education agencies. The education agencies recommended panelists with (a) experience as either superintendents or assistant superintendents, or college faculty who prepare superintendents and (b) familiarity with the knowledge and skills required of beginning superintendents.

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 SSA, the recommended passing score is 96 out of a possible 146 raw-score points. The scale score associated with a raw score of 96 is 162 on a 100–200 scale.

To support the decision-making process for education agencies establishing a passing score (cut score) for the *ETS*[®] School Superintendent Assessment (SSA), research staff from ETS designed and conducted a multistate standard-setting study in December 2018 in Princeton, New Jersey. Education agencies¹ recommended panelists with (a) experience as either superintendents or assistant superintendents, or college faculty who prepare superintendents and (b) familiarity with the knowledge and skills required of beginning superintendents. Eleven states (Table 1) were represented by 18 panelists. (See Appendix A for the names and affiliations of the panelists.)

Table 1
Participating State and Number of Panelists

Idaho (1 panelist)	Pennsylvania (3 panelists)
Kansas (1 panelist)	Rhode Island (1 panelist)
Louisiana (1 panelist)	South Carolina (1 panelist)
Maryland (1 panelist)	West Virginia (3 panelists)
Mississippi (1 panelist)	Wyoming (1 panelist)
New Jersey (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 SSA passing score (see Geisinger & McCormick, 2010). A state may accept the recommended passing score, adjust the score upward to reflect more stringent expectations, or adjust the score downward to reflect more lenient expectations. There is no *correct* decision; the appropriateness of any adjustment may only be evaluated in terms of its meeting the state's needs.

¹ States that currently use any ETS educator licensure test 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 SSA score and the latter, the reliability of panelists' passing-score recommendation. The SEM allows a state to recognize that any test score on any standardized test—including a SSA score—is not perfectly reliable. A test score only *approximates* what a candidate truly knows or truly can do on the test. The SEM, therefore, addresses the question: How close of an approximation is the test score to the *true* score? The SEJ allows a state to gauge the likelihood that the recommended passing score from a particular panel would be similar to the passing scores recommended by other panels of experts similar in composition and experience. The smaller the SEJ, the more likely that another panel would recommend a passing score consistent with the recommended passing score. The larger the SEJ, the less likely the recommended passing score would be reproduced by another panel.

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

OVERVIEW OF THE *ETS*[®] SCHOOL SUPERINTENDENT ASSESSMENT

The *ETS*[®] *School Leadership Series* Study Companion for the School Superintendent Assessment (ETS, in press) describes the purpose and structure of the test. In brief, the test measures the extent to which entry-level superintendents demonstrate the standards-relevant knowledge and skills necessary for competent professional practice. The test is aligned to the National Policy Board for Educational Administration (NPBEA) *Professional Standards for Educational Leaders* (NPBEA, 2015) and the district-level *National Educational Leadership Preparation* (NELP) standards (NPBEA, 2018).

The three-hour assessment contains 120 selected-response items² and three constructed-response items covering seven content areas: *Strategic Leadership* (approximately 26 selected-response items), *Instructional Leadership* (approximately 19 selected-response items), *Climate and Cultural Leadership* (approximately 22 selected-response items), *Ethical and Policy Leadership* (approximately 16 selected-response items), *Organizational Leadership* (approximately 24 selected-response items), *Community Leadership* (approximately 13 selected-response items) and *Integrated Knowledge and Understanding* (3 constructed-response items).³ The reporting scale for the SSA ranges from 100 to 200 scale-score points.

PROCESSES AND METHODS

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

² Ten of the 120 selected-response items are pretest items and do not contribute to a candidate's score.

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

REVIEWING THE TEST

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

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 superintendents or areas that address content particularly important for entry-level superintendents.

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 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 SSA includes both dichotomously-scored (selected-response items) and constructed-response items. Panelists received training in two distinct standard-setting approaches: one standard-setting approach for the dichotomously-scored items and another approach for the constructed-response items.

A panel's passing score is the sum of the interim passing scores recommended by the panelists for (a) the dichotomously-scored items and (b) the constructed-response items. As with scoring and reporting,

the panelists' judgments for the constructed-response items were weighted such that they contributed approximately 25% of the overall score.

Dichotomously scored items. The standard-setting process for the dichotomously-scored items 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. Then the panelists estimated what chance a just qualified candidate would have of answering 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 rationale. All panelists completed a post-training survey to confirm that they had received adequate training and felt prepared to continue; the standard-setting process continued only if all panelists confirmed their readiness.

Constructed-response items. An Extended Angoff method (Cizek & Bunch, 2007; Hambleton & Plake, 1995) was used for the constructed-response items. For this portion of the study, a panelist decided on the assigned score value that would most likely be earned by the just qualified candidate for each constructed-response item. Panelists were asked first to review the description of the just qualified candidate and then to review the constructed-response item and its rubric. The rubric for a constructed-response item defines (holistically) the quality of the evidence that would merit a response earning a particular score. During this review, each panelist independently considered the level of knowledge/skill required to respond to the constructed-response item and the features of a response that would earn a particular score, as defined by the rubric. Each panelist decided on the score most likely to be earned by the just qualified candidate from the possible values a test taker can earn.

A test-taker's response to a constructed-response item is independently scored by two raters, and the sum of the raters' scores is the assigned score⁴; possible scores, therefore, range from zero (both raters assigned a score of zero) to six (both raters assigned a score of three). For their ratings, each panelist decided on the score most likely to be earned by a just qualified candidate from the following possible values: 0, 1, 2, 3, 4, 5, or 6. For each of the constructed-response item, panelists recorded the score (0 through 6) that a just qualified candidate would most likely earn.

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

Multiple Rounds. 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. For dichotomously-scored items, 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

⁴ If the two raters' scores differ by more than one point (non-adjacent), the Chief Reader for that item assigns the score, which is then doubled.

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 18 panelists representing 11 states (See Appendix A for a listing of panelists.) Nine panelists were superintendents, three were assistant (associate or deputy) superintendents, four were college faculty, one was a college administrator, and one was a state coordinator of a principal evaluation system. The job responsibilities of all five panelists working at a college/university included the training of superintendents.

Table 2
Panel Member Demographics

	<i>N</i>	<i>%</i>
Current position		
Superintendent	9	50
Assistant (Associate or Deputy) Superintendent	3	17
College faculty	4	22
College Administrator	1	6
State Coordinator, Principal Evaluation System	1	6
Race		
White	12	67
Black or African American	6	33
Gender		
Female	10	56
Male	8	44
Are you currently certified as a superintendent in your state?		
Yes	12	67
No	0	0
I am not a superintendent	6	33

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

	<i>N</i>	<i>%</i>
Including this year, how many years of experience do you have as a superintendent?		
3 years or less	3	17
4 - 7 years	5	28
8 - 11 years	3	17
12 - 15 years	1	6
16 years or more	0	0
I am not a superintendent	6	33
Which best describes the location of your K-12 school?		
Urban	3	17
Suburban	5	28
Rural	4	22
I am not a superintendent	6	33
Are you currently supervising or mentoring other superintendents or assistant superintendents?		
Yes	4	22
No	8	44
I am not a superintendent	6	33
Are you currently involved in the training or preparation of school superintendents?		
Yes	5	2
No	0	0
I am not college faculty	13	72
How many years of experience (including this year) do you have preparing school superintendents?		
3 years or less	1	6
4 - 7 years	2	11
8 - 11 years	1	6
12 - 15 years	0	0
16 years or more	1	6
I am not college faculty	13	72

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. 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	87.80	87.45
2	86.20	87.90
3	89.50	87.30
4	103.70	103.20
5	99.90	102.10
6	97.30	97.60
7	103.40	103.40
8	100.70	101.00
9	90.80	94.90
10	99.00	96.00
11	108.70	110.80
12	87.80	88.90
13	84.80	91.30
14	79.80	76.90
15	96.50	96.50
16	101.30	97.60
17	98.00	95.90
18	97.15	98.55
Average	95.13	95.41
Lowest	79.80	76.90
Highest	108.70	110.80
SD	7.79	7.85
SEJ	1.84	1.85

The panel’s passing score recommendation for the *ETS*[®] School Superintendent Assessment is 95.41 (out of a possible 146 raw-score points). The value was rounded to the next highest whole number, 96, to determine the functional recommended passing score. The scale score associated with 96 raw points is 162.

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
	96 (5.63)	162
-2 CSEM	85	152
-1 CSEM	91	157
+ 1 CSEM	102	167
+ 2 CSEM	108	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 their 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* 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* 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; 13 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. Two-thirds of the panelists (12 of the 18 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; 14 of the 18 panelists were *very comfortable*. Fifteen of the 18 panelists

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.

indicated the recommended passing score was *about right*; two panelists indicated that the passing score was *too low* and one panelist indicated the passing score was *too high*.

SUMMARY

To support the decision-making process for education agencies establishing a passing score (cut score) for the *ETS*[®] School Superintendent Assessment (SSA), 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 SSA, the recommended passing score is 96 out of a possible 146 raw-score points. The scale score associated with a raw score of 96 is 162 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.
- 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 ETS® School Leadership Series: Study Companion: School Superintendent Assessment (6991)*. 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.
- Hambleton, R. K., & Plake, B.S. (1995). Using an extended Angoff procedure to set standards on complex performance assessments. *Applied Measurement in Education, 8*, 41-55.
- National Policy Board for Educational Administration (2015). *Professional Standards for Educational Leaders 2015*. Reston, VA: Author.
- 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.
- NPBEA. (2018). National Educational Leadership Preparation (NELP) Program Standards - District Level. Retrieved from: www.npbea.org.

APPENDIX A

PANELISTS' NAMES & AFFILIATIONS

Participating Panelists With Affiliation

<u>Panelist</u>	<u>Affiliation</u>
Rich Bauscher	University of Idaho (ID)
Colleen Burns-Jermain	Newport Public Schools (RI)
Bonita Coleman	Ocean Springs School District (MS)
Cory Gibson	Valley Center USD 262 (KS)
Bernadine Habursky	The School District of the City of Erie (PA)
Blaine Hess	Jackson County Schools (WV)
Tom Hisiro	Marshall University (WV)
Kelli Joseph	St. Helena Parich School District (LA)
JoAnn Manning	Rowan University (NJ)
Margaret McLaughlin	University of Maryland-College Park (MD)
Tracy Ragland	Wyoming Professional Teaching Standards Board (WY)
Glenn Robbins	Tabernacle Township School District (NJ)
Tamara Thomas Smith	Cheltenham School District (PA)
Thomas A. Smith	Hopewell Valley Regional School District (NJ)
Vicki Traufler	South Carolina Department of Education (SC)
Tracy Vitale	Seneca Valley School District (PA)
Kevin West	East Orange School District (NJ)
Patricia Wilson	West Virginia State University (WV)

APPENDIX B
STUDY AGENDA

AGENDA

***ETS*[®] School Superintendent Assessment (SSA) Standard-Setting Study**

Day 1

Welcome and Introduction

Overview of Standard Setting and the SSA

Review the SSA

Discuss the SSA

Describe the Knowledge/Skills of a Just Qualified Candidate

Standard Setting Training for Selected-Response Items

Round 1 Judgments for Selected-Response Items

Collect Materials; End of Day 1

Day 2

Overview of Day 2

Standard Setting Training for Constructed-Response Items

Round 1 Judgments for Constructed-Response Items

Round 1 Feedback and Round 2 Judgments

Feedback on Round 2 Recommended Cut Score

Complete Final Evaluation

Collect Materials; End of Study

APPENDIX C

JUST QUALIFIED CANDIDATE DESCRIPTION

Description of the Just Qualified Candidate⁷

A just qualified candidate ...

I. Strategic Leadership

A. Mission, Vision, and Core Values

1. Knows purpose (timeframe, context, etc.) of vision, mission, and strategic plan
2. Understands students first
3. Aware if vision, mission, strategic plan is working
4. Knows how to identify and engage key, diverse stakeholders and gather data
5. Understands the connectiveness of vision, mission, and strategic plan and how to make actionable with leadership team

B. District and School Improvement

1. Understand the importance of strategies for preparing key stakeholders for engaging in district/school improvement
2. Understand conceptual framework for school improvement and how to select strategies to address specific needs
3. Understanding characteristics of credible evidence and how to use/interpret different forms of evidence/data
4. Understand how to create an actionable and measurable school improvement plan

II. Instructional Leadership

A. Curriculum and Instruction

1. Understands the importance of developing, designing, and delivering curriculum that: reflects high expectations, aligns with academic standards, reflects cultural responsiveness
2. Knows the importance of being the advocate for high level, effective pedagogy
3. Understands the importance of the integration of technology across the curriculum to support teaching and enhance learning
4. Understands the role of intervention systems as tools to assist with the identification of students' needs and utilizing data to make informed decisions

B. Assessment and Accountability

1. Values the use of assessment and its importance in measuring and determining student learning
2. Understanding assessment as a means of collecting, analyzing and communicating results with key stakeholders to facilitate informed decisions about continuous improvement
3. Understands the importance of monitoring and evaluating curriculum and instructional practices on a regular basis

⁷ 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 Candidate⁸ (continued)

A just qualified candidate ...

- III. Climate and Culture Leadership
 - A. Equity and Cultural Responsiveness
 - 1. Understand and develop action plans/policies to ensure that students and staff are treated equitably.
 - 2. Understand and serve as an advocate for students to have equitable access to programs, courses and supports that meet the needs of a diverse population.
 - B. Climate and Culture for Staff
 - 1. Understand the importance of work place conditions as they impact student learning.
 - 2. Understand the importance of collective responsibility as a means of supporting a healthy climate and culture
 - 3. Understand how to support a professional culture of engagement for all stakeholders
 - C. Community of Care and Support for Students
 - 1. Can identify characteristics of an inclusive, caring and supporting school district
 - 2. Can identify characteristics of a safe, caring and healthy learning environment in order to meet the needs and engage diverse learners
 - 3. Understand how to provide coherent systems of academic, social, and emotional supports in order to meet the needs of a diverse population.
 - 4. Understands and promotes positive relationships between adults to students and between student to student to promote a respectful environment.
- IV. Ethical and Policy Leadership
 - A. Ethical Practice
 - 1. Understands and models how to act ethically and promote ethical and professional relationships and decision making by administrators and staff
 - 2. Knows how to promote and safeguard the values of an equitable public education
 - B. Policy, Advocacy, and Governance
 - 1. Knows how to interpret, create, and enforce appropriate policies that align with the district's mission and vision and ensure student success
 - 2. Knows the roles, functions, and expectations of the superintendent and district board of education.
 - 3. Knows how to effectively communicate and apply state and federal laws to district policies and procedures (regarding, e.g., board governance, employment, confidentiality of students and staff, rights of students and staff)

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 Candidate⁹ (continued)

A just qualified candidate ...

- V. Organizational Leadership
 - A. Professional Capacity
 - 1. Adheres to local, state and federal laws, regulations, and codes as required for human resource management and consults with appropriate legal staff
 - 2. Promote a culture of continuous learning and improvement to empower administrators and staff
 - 3. Know the importance of professional learning opportunities, recruitment, retention, professional growth, etc.
 - B. Operations and Management
 - 1. Applies appropriate procedures to develop and manage a budget and fiscal resources
 - 2. Follows local, state, and federal laws and regulations for facility management
 - 3. Understands and implements the crisis management plan
 - 4. Developing productive interrelationship skills with various stakeholders
- VI. Community Leadership
 - A. Community Engagement
 - 1. Knows the benefits of effective communications and has awareness that there are a variety of communication methods to engage families and community
 - 2. Is aware of the need to build trust through connecting and partnering with key stakeholders for the benefit of the district and to respond to significant concerns
 - B. Maximizing Community Resources
 - 1. Will be able to identify community resources to promote student learning and school improvement
 - 2. Understands the importance of maintaining a community presence and strategies to access community and alternative resources

⁹ 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.	18	100	0	0	0	0	0	0
• The instructions and explanations provided by the facilitators were clear.	17	94	1	6	0	0	0	0
• The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	18	100	0	0	0	0	0	0
• The explanation of how the recommended cut score is computed was clear.	17	91	1	6	0	0	0	0
• The opportunity for feedback and discussion between rounds was helpful.	18	100	0	0	0	0	0	0
• The process of making the standard-setting judgments was easy to follow.	16	89	2	11	0	0	0	0
• I understood how to use the survey software.	17	94	1	6	0	0	0	0

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	13	72	5	28	0	0		
• The between-round discussions	15	83	3	17	0	0		
• The knowledge/skills required to answer each test question ¹⁰	13	72	4	22	0	0		
• The cut scores of other panel members ¹⁰	7	39	10	56	0	0		
• My own professional experience ¹⁰	12	67	5	28	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 cut score?	14	78	3	17	1	6	0	0
	Too low		About right		Too high			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• Overall, the recommended cut score is:	2	11	15	83	1	6		

¹⁰ One panelist did not respond to this question.