



Listening. Learning. Leading.

Multistate Standard-Setting Technical Report

PRAXIS™ MIDDLE SCHOOL ENGLISH LANGUAGE ARTS (5047)

Licensure and Credentialing Research

ETS

Princeton, New Jersey

March 2013

EXECUTIVE SUMMARY

To support the decision-making process of education agencies establishing a passing score (cut score) for the Praxis™ Middle School English Language Arts (5047) test, research staff from Educational Testing Service (ETS) designed and conducted a multistate standard-setting study.

PARTICIPATING STATES

Panelists from 18 states and Washington, DC were recommended by their respective education agencies. The education agencies recommended panelists with (a) experience as either English teachers or college faculty who prepare English teachers and (b) familiarity with the knowledge and skills required of beginning English 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 Middle School English Language Arts test, the recommended passing score¹ is 81 out of a possible 120 raw-score points. The scaled score associated with a raw score of 81 is 164 on a 100–200 scale.

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

To support the decision-making process for education agencies establishing a passing score (cut score) for the Praxis™ Middle School English Language Arts (5047) test, research staff from ETS designed and conducted a multistate standard-setting study in March 2013 in Princeton, New Jersey. Education agencies² recommended panelists with (a) experience as either English teachers or college faculty who prepare English teachers and (b) familiarity with the knowledge and skills required of beginning English teachers. Eighteen states and Washington, DC (Table 1) were represented by 28 panelists. (See Appendix A for the names and affiliations of the panelists.)

Table 1
Participating Jurisdictions and Number of Panelists

Alaska (2 panelists)	North Carolina (2 panelists)
Arkansas (2 panelists)	North Dakota (1 panelist)
Delaware (1 panelist)	Rhode Island (1 panelist)
Hawaii (1 panelist)	South Dakota (2 panelists)
Kansas (1 panelist)	Utah (2 panelists)
Kentucky (2 panelists)	Vermont (2 panelists)
Louisiana (1 panelist)	Washington, DC (2 panelists)
Maryland (1 panelist)	West Virginia (1 panelist)
New Hampshire (2 panelists)	Wyoming (1 panelist)
Nevada (1 panelist)	

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

ETS provides a recommended passing score from the multistate standard-setting study to education agencies. In each jurisdiction, the department of education, the board of education, or a designated educator licensure board is responsible for establishing the operational passing score in accordance with applicable regulations. This study provides a recommended passing score,³ which represents the combined judgments of two panels of experienced educators. Each jurisdiction may want to consider the recommended passing score but also other sources of information when setting the final

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

³ In addition to the recommended passing score averaged across the two panels, the recommended passing scores for each panel are presented.

Praxis Middle School English Language Arts 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 Middle School English Language Arts test 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 Middle School English Language Arts test score—is not perfectly reliable. A test score only *approximates* what a candidate truly knows or truly can do on the test. The SEM, therefore, addresses the question: How close of an approximation is the test score to the *true* score? The SEJ allows a 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 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.

OVERVIEW OF THE PRAXIS MIDDLE SCHOOL ENGLISH LANGUAGE ARTS TEST

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

The two-hours and forty minutes assessment contains 110 selected-response items⁴ and two constructed-response items covering four content areas: *Reading* (approximately 50 selected-response items and one constructed-response item), *Language Use and Vocabulary* (approximately 16 selected-response items), *Writing, Speaking and Listening* (approximately 26 selected-response items) and *English Language Arts Instruction* (approximately 19 selected-response items and one constructed-response item).⁵ The reporting scale for the Praxis Middle School English Language Arts test ranges from 100 to 200 scaled-score points.

PROCESSES AND METHODS

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

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

REVIEWING THE TEST

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

⁴ Twenty of the 110 selected-response items are pretest items and do not contribute to a candidate's score.

⁵ The number of items for each content area may vary slightly from form to form of the test.

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

DEFINING THE TARGET CANDIDATE

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

Panel 1 created a description of the target candidate — the knowledge/skills that differentiate a *just* from a *not quite* qualified candidate. To create this description, the panel first split into smaller groups to consider the target candidate. The full panel then reconvened and, through whole-group discussion, created the description of the target candidate to use for the remainder of the study.

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

For Panel 2, the panelists began with the description of the target candidate developed by Panel 1. Given that the multistate standard-setting study was designed to provide two recommendations for the same performance standard, it was important that panels use consistent target candidate description to frame their judgments. The panelists reviewed the target candidate description, and any ambiguities were discussed and clarified.

PANELISTS' JUDGMENTS

The Praxis Middle School English Language Arts test 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 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 target candidate would answer the item correctly. Panelists made their judgments using the following rating scale: 0, .05, .10, .20, .30, .40, .50, .60, .70, .80, .90, .95, 1. The lower the value, the less likely it is that the target candidate would answer the item correctly because the item is difficult for the target candidate. The higher the value, the more likely it is that the target candidate would answer the item correctly.

Panelists were asked to approach the judgment process in two stages. First, they reviewed both the description of the target candidate and the item and decided if, overall, the item would be difficult for the target candidate, easy for the target candidate or moderately difficult/easy. The facilitator encouraged the panelists to consider the following rules of thumb to guide their decision:

- Difficult items for the target candidate are in the 0 to .30 range.
- Moderately difficult/easy items for the target candidate are in the .40 to .60 range.
- Easy items for the target candidate are in the .70 to 1 range.

Next, panelists decided how to refine their judgment within the range. For example, if a panelist thought that an item would be easy for the target candidate, the initial decision located the item in the .70 to 1 range. The second decision for the panelist was to decide if the likelihood of answering it correctly is .70, .80, .90, .95 or 1.

After the training, panelists made practice judgments and discussed those judgments and their rationale. All panelists completed a post-training survey to confirm that they had received adequate

training and felt prepared to continue; the standard-setting process continued only if all panelists confirmed their readiness.

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 target candidate for each constructed-response item. Panelists were asked first to review the definition of the target 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 target 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 score⁶; 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 target 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 target 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.

⁶ 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.

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

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

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

RESULTS

EXPERT PANELS

Table 2 presents a summary of the panelists' demographic information. The panel included 28 educators representing 18 states and Washington, DC. (See Appendix A for a listing of panelists.) Twenty panelists were teachers, six were college faculty, one was an administrator or department head, and one held another position. Five of the six faculty members' job responsibilities included the training of English 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)

	<i>N</i>	<i>%</i>
Current position		
Teacher	20	71%
Administrator/Department head	1	4%
College faculty	6	21%
Other	1	4%
Race		
White	23	82%
Black or African American	5	18%
Gender		
Female	23	82%
Male	5	18%
Are you currently certified to teach this subject in your state?		
Yes	26	93%
No	2	7%
Are you currently teaching this subject in your state?		
Yes	21	75%
No	7	25%
Are you currently supervising or mentoring other teachers of this subject?		
Yes	14	50%
No	14	50%
At what K–12 grade level are you currently teaching this subject?		
Middle school (6–8 or 7–9)	19	68%
High school (9–12 or 10–12)	1	4%
Middle and High School	1	4%
Not currently teaching at the K–12 level	7	25%

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

	<i>N</i>	<i>%</i>
Including this year, how many years of experience do you have teaching this subject?		
3 years or less	5	18%
4–7 years	7	25%
8–11 years	7	25%
12–15 years	6	21%
16 years or more	3	11%
Which best describes the location of your K–12 school?		
Urban	4	14%
Suburban	7	25%
Rural	10	36%
Not currently working at the K–12 level	7	25%
If you are college faculty, are you currently involved in the training/preparation of teacher candidates in this subject?		
Yes	5	18%
No	1	4%
Not college faculty	22	79%

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.⁷ 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).

⁷ An SEJ assumes that panelists are randomly selected and that standard-setting judgments are independent. It is seldom the case that panelists are randomly sampled, and only the first round of judgments may be considered independent. The SEJ, therefore, likely underestimates the uncertainty of passing scores (Tannenbaum & Katz, 2013).

Table 3
Summary of Round 2 Standard-setting Judgments

	Panel 1	Panel 2
Average	81.52	78.74
Lowest	73.15	69.65
Highest	91.10	83.70
SD	5.43	4.84
SEJ	1.40	1.34

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 Middle School English Language Arts test are 81.52 for Panel 1 and 78.74 for Panel 2 (out of a possible 120 raw-score points). The values were rounded to the next highest whole number, to determine the functional recommended passing score — 82 for Panel 1 and 79 for Panel 2. The scaled scores associated with 82 and 79 raw points are 165 and 162, 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 Middle School English Language Arts test is 80.13 (out of a possible 120 raw-score points). The value was rounded to 81 (next highest raw score) to determine the functional recommended passing score. The scaled score associated with 81 raw points is 164.

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

Table 4***Passing Scores Within 1 and 2 CSEMs of the Recommended Passing Score⁸***

Recommended passing score (CSEM)		Scale score equivalent
	81 (5.01)	164
-2 CSEMs	71	153
-1 CSEM	76	158
+ 1 CSEM	87	170
+ 2 CSEMs	92	176

Note. CSEM = conditional standard error of measurement.

FINAL EVALUATIONS

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

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

All panelists *strongly agreed* or *agreed* that they understood the purpose of the study. Twenty-three of the 28 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 target candidate was at least *somewhat influential* in guiding their standard-setting judgments; 21 of the 28 panelists indicated the description was *very influential*. All but two of the panelists reported that between-round discussions were at least *somewhat influential* in guiding their judgments. Thirteen of the 28 panelists indicated that their own professional experience was *very influential* in guiding their judgments.

⁸ The unrounded CSEM value is added to or subtracted from the rounded passing-score recommendation. The resulting values are rounded up to the next-highest whole number and the rounded values are converted to scaled scores.

All but one of the panelists indicated they were at least *somewhat comfortable* with the passing score they recommended; 19 of the 28 panelists were *very comfortable*. Twenty-seven of the 28 panelists indicated the recommended passing score was *about right*, the remaining panelist indicated that the passing score was *too high*.

SUMMARY

To support the decision-making process for education agencies establishing a passing score (cut score) for the Praxis Middle School English Language Arts 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 Middle School English Language Arts test, the recommended passing score⁹ is 81 out of a possible 120 raw-score points. The scaled score associated with a raw score of 81 is 164 on a 100–200 scale.

⁹ 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.
- 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: Middle School English Language Arts (5047)*. 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.
- Perie, M. (2008). A guide to understanding and developing performance-level descriptors. *Educational Measurement: Issues and Practice, 27*, 15–29.
- Tannenbaum, R. J., & Katz, I. R. (2013). Standard setting. In K. F. Geisinger (Ed.), *APA handbook of testing and assessment in psychology: Vol. 3. Testing and assessment in school psychology and education* (pp. 455–477). Washington, DC: American Psychological Association.

APPENDIX A

PANELISTS' NAMES & AFFILIATIONS

Participating Panelists With Affiliation

<u>Panelist</u>	<u>Affiliation</u>
Anne Brenner Armstrong	University of Alaska Fairbanks (AK)
Amy Brockway	Olathe Public Schools (KS)
Stephanie Buelow	University of Hawaii at Manoa (HI)
Stephanie Carey	Achorage School District (AK)
Christopher Carter	Davis School District (UT)
Sandra Celauro	Rutland Middle School (VT)
Staci Collins	Northwest Cabarrus Middle School (NC)
Dana Emery	Tooele Junior High School (UT)
Jody Fernandez	Morehead State University (KY)
Adrienne Fortune	Missisquoi Valley Union Middle/High School (VT)
TeKyesha Gault	Conway Public School District (AR)
Katherine M. Golec	Jim Bridger Middle School (NV)
Katesha Harrell	A. G. Cox Middle (NC)
Ronnie Harrison	Southern University (LA)
Lynn Johnson	Trinity Washington University (DC)
Stephanie Kaffenberger	Western Hills Middle School (RI)
Jacob Knodel	Discovery Middle School (ND)
Latwayla Knowlton	Annie Camp Junior High (AR)
Patricia Lamontagne	Pelham Memorial School (NH)
Kelly Neal	Bondurant Middle School (KY)
David W. Nicholson	Stevenson University (MD)

Participating Panelists With Affiliation (continued)

<u>Panelist</u>	<u>Affiliation</u>
Ashley Rousseau	Pine Bluffs Junior/Senior High School (WY)
Evelyn Ruffin-Burris	Bayard Middle School (DE)
Alex Scarelli	Campbell High School (NH)
Brandi Swalve	Aberdeen School District (Holgate Middle School) (SD)
Ashley White	Lenore K-8 School (WV)
Christina Yuknis	Gallaudet University (DC)
Susan Zueger	Sioux Falls School District/Memorial Middle School (SD)

APPENDIX B
STUDY AGENDA

AGENDA

Praxis Middle School English Language Arts (5047) Standard-Setting Study

Day 1

Welcome and Introduction

Overview of Standard Setting and the Praxis Middle School English Language Arts Test

“Take” the Praxis Middle School English Language Arts Test
(Take breaks as needed)

Discuss the Praxis Middle School English Language Arts Test

Define the Knowledge/Skills of a Target Candidate

Lunch

Define the Knowledge/Skills of a Target Candidate (continued)

Break

Standard-Setting Training for Selected-Response Items

Round 1 Standard Setting Judgments for Selected-Response

Collect Materials; End of Day 1

AGENDA

Praxis Middle School English Language Arts (5047) Standard-Setting Study

Day 2

Overview of Day 2

Standard Setting Training for Constructed-Response Questions

Round 1 Standard Setting Judgments for Constructed-Response Questions

Round 1 Feedback & Round 2 Judgments

Break

Round 1 Feedback & Round 2 Judgments (continued)

Lunch

Round 1 Feedback & Round 2 Judgments (continued)

Break

Feedback on Round 2 Recommended Passing Score

Complete Final Evaluation

Collect Materials; End of Study

APPENDIX C

TARGET CANDIDATE DESCRIPTION

Description of the Target Candidate¹⁰

A target candidate ...

Reading

1. Identify and differentiate the defining characteristics of major subgenres
2. Analyze how poetic devices and structure contribute to the meaning of a poem
3. Analyze how differences in characters' POV's, setting and characterization influence the overall meaning and individual elements of a text (mood, tone, conflict, etc.)
4. Understand literal and inferential methods that authors use to convey purpose and perspective within informational texts including organizational pattern, word choice and tone
5. Understand how literal and inferential interpretation of informational text can be supported with textual evidence

Language Use & Vocabulary

6. Understand the functions of syntactical and semantic features (such as affixes) to determine advanced and complex word meaning
7. Has awareness of the dialect and diction across regions, cultural groups and time periods

Writing, Speaking, and Listening

8. Evaluate and assess what constitutes effective writing including strong details, supporting evidence, purpose, format, audience

English Language Arts Instruction

9. Select commonly used research-based approaches to middle grades ELA instruction (reading, writing, speaking, listening and viewing)
10. Knows commonly used research-based approaches to grouping and differentiated instruction to meet specific instructional objectives and to motivate adolescents
11. Understands approaches to and purposes of formative and summative assessment of reading, writing, speaking and listening

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

APPENDIX D

RESULTS

Table D1
Panel Member Demographics (by Panel)

	Panel 1		Panel 2	
	<i>N</i>	%	<i>N</i>	%
Current position				
Teacher	8	53%	12	92%
Administrator/Department head	0	0%	1	8%
College faculty	6	40%	0	0%
Other	1	7%	0	0%
Race				
White	13	87%	10	77%
Black or African American	2	13%	3	23%
Gender				
Female	12	80%	11	85%
Male	3	20%	2	15%
Are you currently certified to teach this subject in your state?				
Yes	14	93%	12	92%
No	1	7%	1	8%
Are you currently teaching this subject in your state?				
Yes	9	60%	12	92%
No	6	40%	1	8%
Are you currently supervising or mentoring other teachers of this subject?				
Yes	8	53%	6	46%
No	7	47%	7	54%
At what K–12 grade level are you currently teaching this subject?				
Middle school (6–8 or 7–9)	9	60%	10	77%
High school (9–12 or 10–12)	0	0%	1	8%
Middle and High School	0	0%	1	8%
Not currently teaching at the K–12 level	6	40%	1	8%

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

	Panel 1		Panel 2	
	<i>N</i>	%	<i>N</i>	%
Including this year, how many years of experience do you have teaching this subject?				
3 years or less	2	13%	3	23%
4–7 years	3	20%	4	31%
8–11 years	6	40%	1	8%
12–15 years	3	20%	3	23%
16 years or more	1	7%	2	15%
Which best describes the location of your K–12 school?				
Urban	1	7%	3	23%
Suburban	5	33%	2	15%
Rural	3	20%	7	54%
Not currently working at the K–12 level	6	40%	1	8%
If you are college faculty, are you currently involved in the training/preparation of teacher candidates in this subject?				
Yes	5	33%	0	0%
No	1	7%	0	0%
Not college faculty	9	60%	13	100%

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

	Panel 1		Panel 2	
Panelist	Round 1	Round 2	Round 1	Round 2
1	72.05	80.80	77.50	81.35
2	77.40	75.50	84.55	83.35
3	90.85	91.10	80.65	80.85
4	87.80	82.85	74.65	73.80
5	86.35	86.75	68.95	77.75
6	78.05	77.45	86.95	83.70
7	87.80	84.60	79.40	80.30
8	85.00	82.10	64.65	69.75
9	72.85	73.75	65.80	69.65
10	72.70	76.60	85.90	82.90
11	71.75	73.15	78.80	79.10
12	95.70	88.55	81.60	82.70
13	77.75	80.25	78.05	78.40
14	82.50	85.85		
15	73.85	83.45		
Average	80.83	81.52	77.50	78.74
Lowest	71.75	73.15	64.65	69.65
Highest	95.70	91.10	86.95	83.70
SD	7.73	5.43	7.22	4.84
SEJ	2.00	1.40	2.00	1.34

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

	Strongly agree		Agree		Disagree		Strongly disagree	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
• I understood the purpose of this study.	14	93%	1	7%	0	0%	0	0%
• The instructions and explanations provided by the facilitators were clear.	11	73%	4	27%	0	0%	0	0%
• The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	11	73%	4	27%	0	0%	0	0%
• The explanation of how the recommended passing score is computed was clear.	13	87%	2	13%	0	0%	0	0%
• The opportunity for feedback and discussion between rounds was helpful.	12	80%	3	20%	0	0%	0	0%
• The process of making the standard-setting judgments was easy to follow.	9	60%	6	40%	0	0%	0	0%

Table D3 (continued)
Final Evaluation: Panel 1

How influential was each of the following factors in guiding your standard-setting judgments?	Very influential		Somewhat influential		Not influential			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• The description of the target candidate	9	60%	6	40%	0	0%		
• The between-round discussions	7	47%	6	40%	2	13%		
• The knowledge/skills required to answer each test item	13	87%	2	13%	0	0%		
• The passing scores of other panel members	3	20%	8	53%	4	27%		
• My own professional experience	5	33%	9	60%	1	7%		
	Very comfortable		Somewhat comfortable		Somewhat uncomfortable		Very uncomfortable	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
• Overall, how comfortable are you with the panel's recommended passing score?	9	60%	5	33%	1	7%	0	0%
	Too low		About right		Too high			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• Overall, the recommended passing score is:	0	0%	14	93%	1	7%		

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

	Strongly agree		Agree		Disagree		Strongly disagree	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
• I understood the purpose of this study.	13	100%	0	0%	0	0%	0	0%
• The instructions and explanations provided by the facilitators were clear.	12	92%	1	8%	0	0%	0	0%
• The training in the standard-setting method was adequate to give me the information I needed to complete my assignment.	11	85%	2	15%	0	0%	0	0%
• The explanation of how the recommended passing score is computed was clear.	9	69%	4	31%	0	0%	0	0%
• The opportunity for feedback and discussion between rounds was helpful.	9	69%	4	31%	0	0%	0	0%
• The process of making the standard-setting judgments was easy to follow.	9	69%	4	31%	0	0%	0	0%

Table D4 (continued)
Final Evaluation: Panel 2

How influential was each of the following factors in guiding your standard-setting judgments?	Very influential		Somewhat influential		Not influential			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• The description of the target candidate	12	92%	1	8%	0	0%		
• The between-round discussions	6	46%	7	54%	0	0%		
• The knowledge/skills required to answer each test item	11	85%	2	15%	0	0%		
• The passing scores of other panel members	1	8%	9	69%	3	23%		
• My own professional experience	8	62%	3	23%	2	15%		
	Very comfortable		Somewhat comfortable		Somewhat uncomfortable		Very uncomfortable	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
• Overall, how comfortable are you with the panel's recommended passing score?	10	77%	3	23%	0	0%	0	0%
	Too low		About right		Too high			
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%		
• Overall, the recommended passing score is:	0	0%	13	100%	0	0%		