

**Standard Setting in Vermont
for Selected Tests in The Praxis Series™
Study Date: May 2009**

**Prepared by:
Educational Testing Service
July 2009**



Prepared for the Vermont Department of Education

Copyright © 2009 by Educational Testing Service. All rights reserved.

EDUCATIONAL TESTING SERVICE, ETS, the modernized ETS logo, PPST, THE PRAXIS SERIES: PROFESSIONAL ASSESSMENTS FOR BEGINNING TEACHERS and its logo are registered trademarks of Educational Testing Service. THE PRAXIS SERIES is a trademark of Educational Testing Service.

Educational Testing Service is an Equal Opportunity Employer.

STANDARD SETTING IN VERMONT FOR SELECTED TESTS IN THE PRAXIS SERIES™

INTRODUCTION

A standard setting study was conducted by Educational Testing Service (ETS), in conjunction with the Vermont Department of Education (VTDOE), to set standards for two tests in The Praxis Series™:

- English to Speakers of Other Languages (0360)
- Reading Specialist (0300)

The study was conducted on May 27, 2009, involving a total of 30 Vermont educators.

The primary purpose of the study was to provide VTDOE with recommendations regarding a passing score or cutscore for each test. Each panel also collected content-related validity evidence of the importance of the content specifications for entry-level teachers in Vermont.

Scores for the two tests are reported on a scale that ranges from 250-990.

PANEL DEMOGRAPHICS

Panelists were selected by the VTDOE based on criteria designed to help ensure that panelists (1) were familiar with the job requirements in the relevant content area and the capabilities of a beginning teacher, (2) were representative of Vermont public school educators in terms of gender, race/ethnicity, and geographic region within the state, and (3) consisted primarily of certified teachers with 3 to 10 years of experience. Table 1 summarizes the demographic characteristics of the panelists. Highlights from the table are presented below.

- **English to Speakers of Other Languages (0360)**
All of the panelists were English as a Second Language teachers. Twelve panelists are certified to teach and 11 are currently teaching English as a Second Language. Nine panelists had between 4 and 11 years of experience and three had at least 12 years of experience. Six panelists reported mentoring another teacher. Twelve of the panelists were female, and 12 were White.
- **Reading Specialist (0300)**
Sixteen of the panelists were teachers and one was college faculty member. Fifteen of the panelists are certified and 14 currently reading specialists. One panelist had 3 years or less experience, 10 of the panelists had between 4 and 11 years of experience and six had at least 12 years of experience. Three panelists reported mentoring another reading specialist. Sixteen of the panelists were female, and all of the panelists were White.

TABLE 1
SUMMARY OF PANEL DEMOGRAPHICS

	English to Speakers of Other Languages (N = 13)		Reading Specialist (N = 17)	
	N	Percent	N	Percent
Group you are representing				
Teachers	13	100%	16	94%
College Faculty	0	0%	1	6%
Race				
White	12	92%	17	100%
Hispanic	1	8%	0	0%
Gender				
Female	12	92%	16	94%
Male	1	8%	1	6%
Are you certified by Vermont to teach this subject?				
No	0	0%	1	6%
Yes	12	92%	15	88%
No Response	1	8%	1	6%
Are you currently teaching this subject?				
No	1	8%	2	12%
Yes	11	85%	14	82%
No Response	1	8%	1	6%
Are you currently mentoring another teacher in this subject?				
No	6	46%	12	71%
Yes	6	46%	3	18%
No Response	1	8%	2	12%
How many years of experience do you have teaching this subject in Vermont?				
3 years or less	0	0%	1	6%
4 – 7 years	5	38%	2	12%
8 – 11 years	4	31%	8	47%
12 – 15 years	1	8%	0	0%
16 years or more	2	15%	6	35%
No Response	1	8%	0	0%
What educational level are you currently teaching?				
K - Grade 3	0	0%	2	12%
K – Grade 6	0	0%	1	6%
K – Grade 8 (Elementary)	5	38%	7	41%
Grades 1-4	0	0%	1	6%
Grades 5 – 8 (Middle Level)	1	8%	1	6%
Grades 7 -12 (Secondary)	6	46%	4	24%
Grades 9-12	1	8%	0	0%
Pre K – 12 Literacy Consultant	0	0%	1	6%

TABLE 1 (continued)
SUMMARY OF PANEL DEMOGRAPHICS

	English to Speakers of Other Languages (N = 13)		Reading Specialist (N = 17)	
	N	Percent	N	Percent
School Setting				
Urban	7	54%	1	6%
Suburban	2	15%	7	41%
Rural	4	31%	9	53%

PROCESS AND METHOD

The standard setting study began with introductions and a review of the purpose of the standard setting meeting. The facilitators discussed the importance of maintaining test security and asked the panelists to complete confidentiality/non-disclosure and biographical information forms (see the Appendix for copies). Linda Hendrickson, from the VTDOE, welcomed the panelists and provided an overview of the certifications in Vermont for each panel.

The first major standard setting activity for each panel was for the panelists to “take the test” for which they were recommending a cut score (passing score). The purpose of “taking the test” was to enable the panelists to become familiar with the test content. After “taking the test,” each member was given the key (correct answers for the multiple-choice questions) to self-score; how well a member did on the test was not shared. Then, all of the panelists engaged in a discussion of the test content. The room facilitator encouraged the panelists to discuss, for example, what the test was covering (the major content areas), whether they thought any questions would be particularly challenging for beginning teachers in the content area, and whether there were any questions that they thought addressed knowledge or skills that were particularly important for beginning teachers to know or have. The panelists also had the opportunity to raise and discuss concerns about the questions, although it was made clear that the primary purpose of this test familiarization activity was not to conduct a question review, per se. To maintain focus on the standard setting process, but not to discount the panelists’ reactions to questions, they were encouraged to write down specific question concerns, which would then be shared with ETS test development staff.

Following this discussion, each panel defined the Just Qualified Candidate (JQC) for its particular content area. The JQC is the test taker who has the minimum level of knowledge and/or skills believed necessary in the content area addressed by the test to be considered qualified to enter the teaching profession in Vermont. The JQC definition is the operational definition of the cutscore. The goal of the standard-setting process is to identify the test score that aligns with this definition of the JQC. For each panel, the panelists were split into two groups and each group was asked to write down its definition of a JQC. Each group was asked to refer to the National disciplinary content standards, consider the content addressed by the test, and draw upon their own professional experience to guide their definition. Each group posted its definition on chart paper and a full-panel discussion occurred to reach consensus on a final definition for each content area (see Appendix).

Standard Setting

Standard Setting for Multiple-Choice Questions. A probability-based Angoff method (Brandon, 2004; Hambleton & Pitoniak, 2006) was used for multiple-choice questions. In this approach, for each question, a panelist decides on the likelihood (probability or chance) that a JQC would answer it 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 a JQC would answer the question correctly, because the question is difficult for the JQC. The higher the value, the more likely it is that a JQC would answer the question correctly. The panelists were asked to approach the judgment process in

two stages. First, they reviewed the definition of the JQC and the question and decided if, overall, the question was difficult for the JQC, easy for the JQC, or moderately difficult/easy. The facilitator encouraged the panelists to consider the following rule of thumb to guide their decision:

- difficult questions for a JQC were in the 0 to .30 range;
- easy questions were in the .70 to 1 range; and
- moderately difficult/easy questions were in the .40 to .60 range.

The second decision was for panelists to decide how they wanted to refine their judgment within the range. For example, if a panelist thought that a question was easy for a JQC, the initial decision located the question in the .70 to 1 range. The second decision was for the panelist to decide if the likelihood of answering it correctly was .70, .80, .90, .95, or 1.0. The two-stage decision-process was implemented to reduce the cognitive load placed on the panelists.

The panelists then were given an opportunity to practice making their standard setting judgments on 10 multiple-choice questions from a different form of the test. The judgments were collected, summarized, and discussed. The judgments were summarized in terms of difficulty ranges (0 to .30, .40 to .60, and .70 to 1). This clustering was a useful way of showing how similar the judgments were in relation to the overall judged difficulty of each question; that is, whether the majority of panelists saw the same question as being in the same range of difficulty or if the panelists' judgments were more spread out across the three ranges of difficulty. The P+ value, which is the percentage of candidates who answered the question correctly, was given to show the difficulty of each test question. The average of the panel was also presented. The panelists were asked to share their judgment rationales so that others could understand their perspectives. The discussion focused on the connection between the knowledge and skills measured by a question and the knowledge and skills of JQCs.

Upon completion of the practice exercise, panelists completed a readiness-to-proceed form (see the Appendix for a copy). The form asked the panelists to indicate the extent to which they understood the purpose of the standard setting study, the extent to which explanations were clear, and whether or not they were ready to proceed with their actual standard setting judgments.

Once the facilitator verified that all panelists were ready to proceed, the panel began to make their standard setting judgments. For each test question, each panelist judged what is the probability (e.g., 0, .05, .1, . . . 1) that a JQC would answer the question correctly.

Judgment of Test Content Specifications

Following the standard setting process, panelists judged the importance of the knowledge and/or skills stated in or implied by the test content specifications for the job of an entry-level teacher for the test they were reviewing in Vermont. These judgments addressed the perceived content-based validity of each test. Judgments were made using a four-point Likert scale — *very important*, *important*, *slightly important*, and *not important*. The panel first judged the importance of the first content category and its sub-categories. As a group, the panel discussed their judgments and were allowed to revise their judgments following the discussion. The panel independently judged the remaining categories and sub-categories.

RESULTS

Readiness to Proceed Forms

The panelists completed a readiness to proceed form, after they were trained in how to make their standard-setting judgments for the multiple-choice questions. The primary information collected from these forms was the panelists indicating if they had received adequate training to make their standard-setting judgments and were ready to proceed. On each panel, all panelists indicated that they were prepared to make their judgments.

Summary of Standard Setting Judgments

Table 2 presents a summary of the standard-setting judgments for the two multiple-choice tests. The numbers in the table reflect the recommended cutscore—the number of raw points needed to “pass” the test—of each panelist. The panel's average recommended cutscore and highest and lowest cutscores are

reported, as are the standard deviation (SD) of panelists' cutscores and the standard error of judgment (SEJ). The SEJ is one way of estimating the reliability of the judgments. It indicates how likely it would be for other panels of educators similar in make-up, experience, and standard-setting training to the current panel to reach the same cutscore. A comparable panel's cutscore would be within 1 SEJ of the current average cutscore 68 percent of the time and within 2 SEJs 95 percent of the time.

The panels' recommended cut scores and the recommended cutscores plus/minus 1 and 2 Standard Errors of Measurement (SEM) are reported in Table 3, as are the number of points available on the test, percent of maximum (the raw study value divided by the number of total points available), and SEM. The recommended scaled cut scores are:

- English to Speakers of Other Languages = 570
- Reading Specialist = 550

TABLE 2
SUMMARY OF STANDARD SETTING JUDGMENTS

Panelist	0360 English to Speakers of Other Languages (4CNT1)	0300 Reading Specialist (4CNT2)
	Cutscore	Cutscore
1	68.55	65.35
2	74.15	69.50
3	67.95	84.55
4	63.55	44.60
5	74.60	57.50
6	67.80	57.25
7	59.50	66.50
8	74.65	75.50
9	56.80	82.55
10	62.85	64.10
11	81.75	80.80
12	77.00	80.85
13	82.40	69.20
14		73.05
15		81.60
16		50.10
17		86.25
Average	70.12	69.96
SD	8.11	12.43
SEJ	2.25	3.01
Highest	82.40	86.25
Lowest	56.80	44.60

TABLE 3
STANDARD SETTING JUDGMENTS RESULTS

	0360		0300	
	English to Speakers of Other Languages (4CNT1)		Reading Specialist (4CNT2)	
	Recommended Cutscore	Scale Score Equivalent	Recommended Cutscore	Scale Score Equivalent
	71	570	70	550
# of Points	120		114	
% of Maximum	59%		61%	
SEM	4.60		4.61	
- 2 SEM	63	520	62	500
-1 SEM	67	550	66	530
+1 SEM	76	610	75	580
+ 2 SEM	81	640	80	610

Notes: Consistent with the recommended cutscore, the cutscores at the different SEMs have been rounded to the next highest whole number.
Six of the questions on the Reading Specialist test were not scored.

Summary of Specification Judgments

Panelists judged the extent to which the knowledge and/or skills reflected by the content specifications for each test were important for entry-level teachers in Vermont. Panelists rated the content categories and their accompanying sub-categories, on a four-point scale ranging from very important to not important. The panelists' ratings are summarized in Tables 4 and 5.

- English to Speakers of Other Languages (0360)**
Overall, 100 percent of the panelists who judged the three content categories agreed that they were “very important” or “important” for entry-level English as a Second Language teachers. In particular, panelists judged Pedagogy to be “very important.” A majority of panelists (80% or greater) also judged four of the five sub-categories in content category I and all of the sub-categories in content categories II and III to be “very important” or “important.”
- Reading Specialist (0300)**
Overall, 100 percent of the panelists judged three of the four content categories to be either “very important” or “important” for entry-level Reading Specialists. One of the 17 panelists believed that the Reading Leadership content category was only “slightly important.” A majority of panelists (80% or greater) also judged all the sub-categories to be “very important” or “important.”

**English as a Second Language
Just Qualified Candidate**

- Demonstrates understanding of current SLA theory
- Demonstrates understanding of linguistic structures of the English language
- Applies knowledge of how linguistic and cultural diversity affects Language Learning
- Plans and implements effective ELL screening, instruction, accommodations, and assessments in all four domains
- Identifies federal and state legal requirements and guidelines relevant to ELL students

**Reading Specialist/Coordinator
Just Qualified Candidate**

- Working knowledge of the terminology related to writing, word study and the five elements of reading, as well as instructional best practice and assessment
- Demonstrates fundamental understanding of developmental continuum of literacy acquisition and the structure of the English language
- Competent analysis and interpretation of assessment data to plan appropriate instruction for students at the level of the individual student, or for groups of students at the classroom, grade-level, school or district levels
- Functional knowledge of assessments and their purposes
- Functional understanding of effective models of professional development, collaboration, and communication with a variety of stakeholders
- Basic understanding of how to select high-quality instructional materials (e.g., informational texts, literature, technology, resources, etc.)

VERMONT ENGLISH AS A SECOND LANGUAGE PANEL

Number of Panelists = 13

May 27, 2009

Susan Blethen	Burlington High School/Burlington School District
Deborah L. Coombs	Windham Southeast Supervisory Union
Jennifer Course	Brattleboro Union High School/ Windham Southeast Supervisory Union
Sara Crothers	Burlington High School/Burlington School District
Barbara K. Dall	Barre City Elementary and Middle School/Barre Supervisory Union 61
Sylvia Fagin	Main Street Middle School/Montpelier Public Schools
Ruth Ann Hicks	Barre City Elementary and Middle School/Barre Supervisory Union 61
Victoria R. Hingston	Springfield School District
Marianne McNamara	Montpelier High School/Montpelier Public Schools
Ana Rawson	Brattleboro Union High School/Windham Southeast Supervisory Union
Robin Rowell	Bradford Elementary School/ Orange East Supervisory Union
Ingeborg White	Winooski High School/Winooski School District
Solange Zwicker	Marion Cross Elementary School/Norwich School District

VERMONT READING SPECIALIST/COORDINATOR PANEL

Number of Panelists = 17

May 27, 2009

Jeanne Avicolti	Marion Cross School
Nancy H. Behun	Hinesburg Community School/Chittenden South Supervisory Union
Reid A. Burris	White River School/Hartford School District
Gail C. Chase	Albert Lawton Intermediate School/Chittenden Central Supervisory Union
Diane R. Cote	Charlotte Central School/Chittenden South Supervisory Union
Marion J. Dewey	Flood Brook Union School/Windsor Southwest
Pamela Fisher-Parro	Hardwick Elementary School/Orleans Southwest Supervisory Union
Linda L. Francis	Shoreham Elementary School/Addison Central Supervisory Union
Kristin M. Gehsmann	Saint Michael's College
Valerie Kelty	Oxbow High School/Orange-East Supervisory Union
Barbara Kohn-Saxe	Albert D. Lawton Intermediate School/Essex Junction
Jennifer MacHarg	East Montpelier Elementary School/Washington Central Supervisory Union
Claire Reczek	Mettawee Community School/Bennington-Rutland Supervisory Union
Christine Reighley	Sheldon Elementary School/Franklin Northwest Supervisory Union
Linda Weber	Mount Anthony Union High School/Southwest Vermont Supervisory Union
Charlene Webster	Mount Anthony Union High School/Southwest Vermont Supervisory Union
Eileen Zazzaro	Mount Anthony Union High School/ Southwest Vermont Supervisory Union

Test Information - 2007-2008 Total Examinees

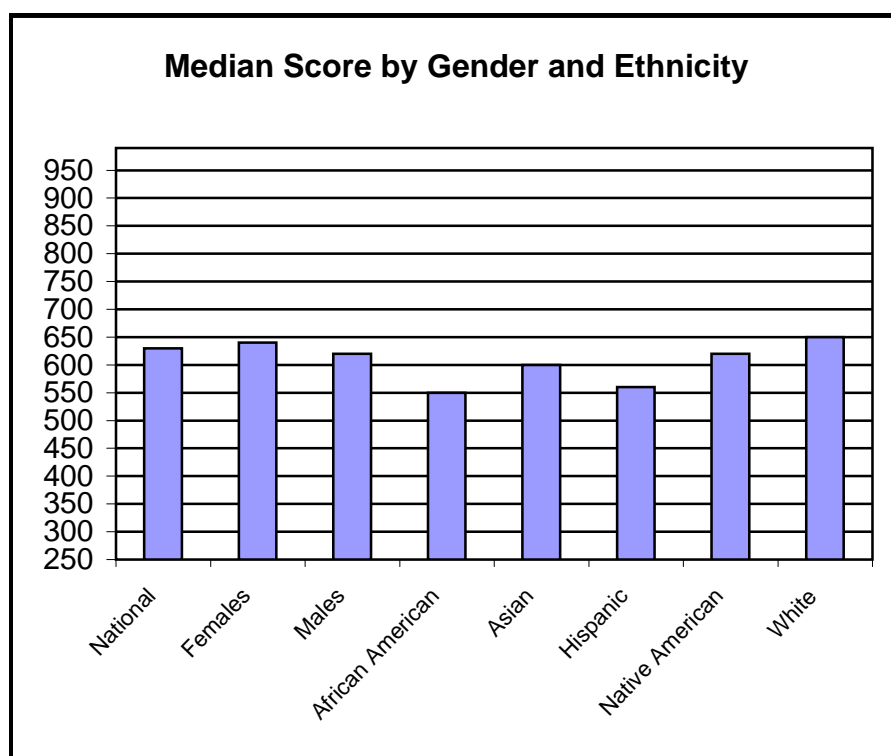
September 1, 2007-August 31, 2008

TEST CODE	0360	TEST NAME	English to Speakers of Other Languages
------------------	-------------	------------------	---

COUNTS	GENDER		ETHNICITY				
Total	Females	Males	African American	Asian	Hispanic	Native American	White
3,733	3,228	502	155	225	373	21	2,632

MEDIAN SCORE	
National	630
Females	640
Males	620
African American	550
Asian	600
Hispanic	560
Native American	620
White	650

<i>Scores in Effect as of 11/1/08</i>	
STATE	PASSING SCORE
KY	620
MN	600
ID	580
WA	580
MD	570
VI	570
WY	560
AL	540
ME	540
SC	540
TN	530
WI	530
DC	520
NC	520
OR	520
HI	510
KS	500
OH	420

**NOTES:**

State passing scores listed in the table do not include states that use multiple passing scores, multiple test composite scores or states that have no minimum passing score.

A state and score listed in *Italics* indicates a score used for Highly Qualified Status Only.

Test Information - 2007-2008 Total Examinees

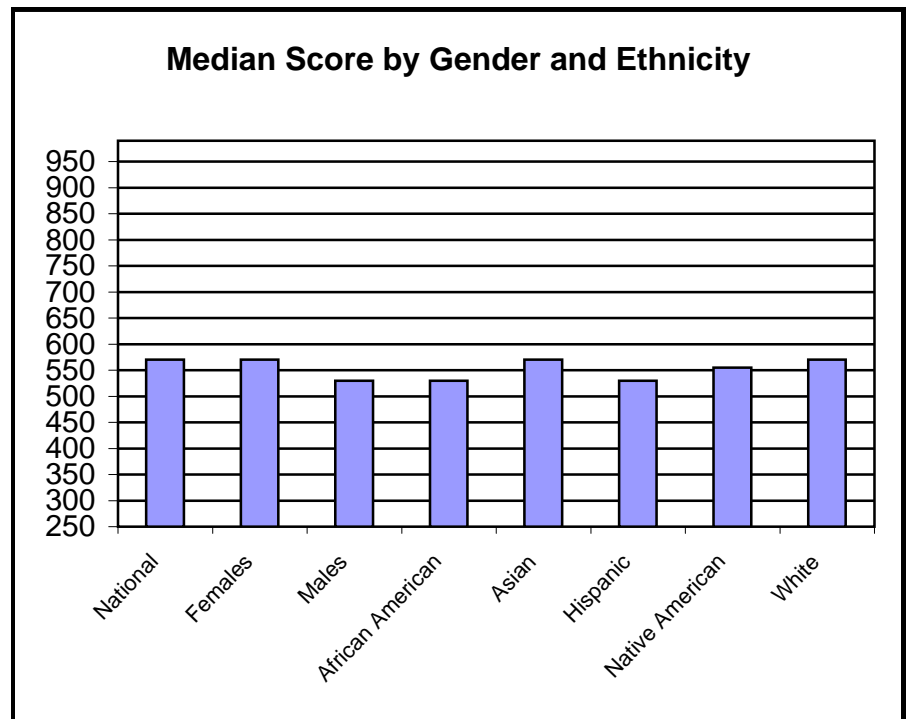
September 1, 2007-August 31, 2008

TEST CODE	0300	TEST NAME	Reading Specialist
-----------	------	-----------	--------------------

COUNTS	GENDER		ETHNICITY				
Total	Females	Males	African American	Asian	Hispanic	Native American	White
5,725	5,251	471	183	44	85	24	5,125

MEDIAN SCORE	
National	570
Females	570
Males	530
African American	530
Asian	570
Hispanic	530
Native American	555
White	570

Scores in Effect as of 11/1/08	
STATE	PASSING SCORE
OR	610
MN	590
NC	570
PA	570
AR	560
DE	560
KS	560
HI	540
WA	540
AL	530
ME	530
WV	520
TN	510
ID	480

**NOTES:**

State passing scores listed in the table do not include states that use multiple passing scores, multiple test composite scores or states that have no minimum passing score.

A state and score listed in *Italics* indicates a score used for Highly Qualified Status Only.