4B

Action

Professional Services Committee

Adoption of the Passing Score Standard for the California Subject Examinations for Teachers (CSET): Mathematics Examination

Executive Summary: This agenda item provides (a) foundational information about the standard setting process and (b) recommendations for passing score standards for the CSET: Mathematics examination, which has been updated to align with the Common Core State Standards.

Policy Question: Does the recommended passing score standard for the CSET: Mathematics examination meet Commission expectations?

Recommended Action: That the Commission adopt the recommended passing score standard for the CSET: Mathematics examination.

Presenter: Michael Taylor, Consultant, Professional Services Division

Strategic Plan Goal

I. Educator Quality

b) Develop, maintain, and promote high quality authentic, consistent educator assessments and examinations that support development and certification of educators who have demonstrated the capacity to be effective practitioners.

Adoption of the Passing Score Standard for the California Subject Examinations for Teachers (CSET): Mathematics Examination

Introduction

This report describes the standard setting study for the California Subject Examination for Teachers Mathematics (CSET): Mathematics Examination, and provides recommended initial passing standards for each subtest based on the recommendations from the CSET: Mathematics Standard Setting Panels.

Based on established examination development procedures, which are outlined in the August 2009 agenda item "Examinations Development Procedures and State Contracting Processes" available at http://www.ctc.ca.gov/commission/agendas/2009-08/2009-08-2D.pdf, the CSET: Mathematics Subject Matter Requirements (SMRs) were revised to align with California's Common Core State Standards (CCSS). The current item bank was reviewed against the CCSS, with appropriate modifications and/or deletions, and some new CCSS-specific test questions were developed. These two phases are discussed in the Background section of this agenda item.

The procedures used in the final phase, setting the passing score, are detailed in the "CSET: Mathematics Standard Setting Studies" section of this item. The CSET: Mathematics Standard Setting Panel and staff recommendations on passing score standards are also provided.

Background: The Standard Setting Process

"Standard setting" is the common term used in the large-scale assessment industry to describe the process of determining a minimum passing score, or cut score, for new or revised examinations. The term "standard" as it is used in standard setting refers to a performance standard, or minimum level of acceptable performance on an examination.

For criterion-referenced examinations like the CSET, standard setting is a content-focused, structured process in which a panel of content area experts reviews the content of an examination, carefully considers the knowledge and skills being measured and relevant data such as question difficulty levels and potential pass rates for various cut scores to make an informed judgment about the minimum level of content knowledge that examinees should demonstrate to "pass" the examination. The standard setting process results in a recommended cut score from the content expert panel to the Commission, which has the authority to establish a minimum passing standard for CSET.

Standard setting is a common and established process for determining valid and defensible minimum passing scores for examinations. Standard setting allows an authoritative body, in this

case the Commission, to make an informed decision when establishing cut scores instead of arbitrarily selecting a minimum passing standard.

There have been many different methods for standard setting published and researched in the field of large-scale assessment over the last 50 years. There are many different standard setting methods in use today for various types of assessments all over the world. All of the most common standard setting methods for educational assessments involve the informed judgments of "raters," or content area experts.

For Commission-owned examinations raters are recruited and selected by an application process which is managed through the recruitment website established by the Commission's contractor specifically to recruit California educators to serve on expert panels for various tasks associated with development and administration of Commission-owned examinations. The recruitment website can be found at http://www.carecruit.nesinc.com/.

The specific standard setting process used for CSET: Mathematics is described more fully below.

Background: Subject Matter Competence Requirement for Candidates

Candidates for a California Multiple or Single Subject preliminary teaching credential have to demonstrate subject matter competence as one of the requirements for the credential. In accordance with the State Board of Education's Highly Qualified Teacher compliance plan, Single Subject candidates may meet the subject matter competence requirement by completing a Commission-approved subject matter preparation program or by passing the applicable CSET Single Subject examination.

Underlying both of these routes to verify subject matter knowledge is a common set of subject matter requirements (SMRs). The SMRs define the content that is eligible to be included on the subject matter examination as well as what must be covered within an approved subject matter program's coursework.

SMRs serve multiple purposes and functions:

- Structuring test content to be clear and understandable to professionals in the field and candidates preparing for the assessment
- Providing meaningful categories for test design and the development of test items
- Informing the general public, legislators, and other constituencies about test content and expectations for public school teachers
- Supporting the use of consistent scoring criteria and procedures
- Providing a framework for reporting test scores to candidates, preparation programs, the public, and the Commission
- Providing a framework for subject matter preparation program standards to assure that candidates who complete the examination route to demonstrating subject matter competence and those who complete the subject matter program route have the same underlying content knowledge

The revised SMRs for CSET: Mathematics were adopted by the Commission in June 2013 (http://www.ctc.ca.gov/commission/agendas/2013-06/2013-06-4C.pdf). The SMRs for CSET: Mathematics are outlined in Appendix A.

Development of the Revised and New CSET: Mathematics Test Items

The CSET contractor, the Evaluation Systems group of Pearson, worked to modify and/or develop a full range of test items to appropriately measure the revised subject matter requirements. Existing items in the test bank were reviewed by the Commission's standing Bias Review Committee for appropriateness and lack of bias, and then by the appointed Content Expert review panels for content focus and consistency with the SMRs. These items were revised where necessary. In addition, new items were developed to fill any existing "gaps" in coverage of the revised SMRs. These new items were then reviewed by both the Bias Review Committee and the content expert panels for each content area.

CSET: Mathematics Test Structure

The SMRs define the content measured by the revised CSET subject matter examinations. The CSET: Mathematics examination is comprised of three subtests which together as a whole address all of the SMRs. The structure of the revised examination is shown in the table below.

Table 1: CSET: Mathematics Test Structure

CSET: Mathematics			
Sub-test	**Domains**	**Number of Multiple-Choice Questions**	**Number of Constructed-Response Questions (short [focused] responses)**
I	Number and Quantity Algebra	10 25	1 2
	Subtest Total	**35**	**3**
II	Geometry Probability and Statistics	25 10	2 1
	Subtest Total	**35**	**3**
III	Calculus	30	2
	Subtest Total	**30**	**2**
	TOTAL	**100**	**8**

Responses to multiple-choice questions are machine scored as correct or incorrect. There is no penalty for guessing. Responses to the constructed-response assignments are scored independently by at least two qualified and well-trained California educators using standardized procedures. Responses are scored using a three-point score scale for focused constructed-response questions and a four-point score scale for extended constructed-response items.

Historical Information about the Previous Version of CSET: Mathematics

Table 2 below shows the passing rates for each year of the previous version of the exam. The CSET: Mathematics examination was first administered in 2002. Passing rates are shown for examinees' first attempt at passing all three subtests (first attempt) and separately for examinees' most successful attempt at passing all three subtests (best attempt) for each year the exam was administered (2002- 2014).

Table 2: CSET: Mathematics Passing Rates (all three subtests) 2002-2014

Program Year	First Attempt Pass Rate	Best Attempt Pass Rate
All Years	**35%**	**64%**
2002-2003	18%	49%
2003-2004	32%	63%
2004-2005	36%	65%
2005-2006	34%	67%
2006-2007	37%	69%
2007-2008	35%	65%
2008-2009	41%	68%
2009-2010	42%	70%
2010-2011	44%	71%
2011-2012	39%	61%
2012-2013	58%	77%
2013-2014	61%	72%

Passing rates for the previous version of CSET: Mathematics started low and generally increased over time during the life of the examination. This phenomenon is common with new assessments and is generally understood to be the result of examinees and programs, where applicable, becoming more familiar with the content being measured by the examination.

The CSET: Mathematics Standard Setting Studies

The purpose of standard setting studies is to provide the Commission with recommendations, based on the informed judgments of California educators, relevant to the determination of the initial passing standards for the CSET: Mathematics examination. The educators on the Standard Setting Panels represented teachers in the content area, district-level educators, and teacher preparation program faculty responsible for the preparation of multiple subject teachers and/or and single subject English teachers. Because of the timing of the standard

setting study relative to agenda item deadlines, demographic information about the specific educators who served on the standard setting panels will be provided in the In-folder item.

As with the standard setting study method used for all other Commission examinations, the process employed for the CSET: Mathematics exam was consistent with recognized psychometric principles and procedures. The standard setting study for CSET: Mathematics was conducted on February 3, 2015.

The CSET: Mathematics standard setting meeting began with an orientation and training session. The initial step was to ask the panel members to independently take the examination under simulated test-like conditions. This helped the members become familiar with the examination, the knowledge and skills associated with the items, and the perspective of the examinees. The panel members were then familiarized with the SMRs and the concept of the minimally competent level of subject knowledge for a beginning teacher. Panel members were asked to conceptualize the specific content knowledge and skills of a hypothetical teacher candidate who would be minimally competent in the subject area. Panel members used this concept of what a minimally competent new teacher would know and be able to do in determining their recommended minimally acceptable score for passing each subtest. Although a number of examinees may exceed the level of acceptable knowledge and skills, none receiving a passing score should fall below this minimally competent level. The panel also reviewed the performance characteristics and score scales used to evaluate the constructed-response items in the CSET: Mathematics examination, which are provided in Appendix B. After this extensive training and the simulated test taking, panel members completed the following three rounds of standard setting activities, as described below, that focused on arriving at an informed judgment as to what the potential cut score should be that reflects the minimum level of subject matter knowledge necessary for a beginning practitioner just competent to begin professional practice.

- Round One: For each multiple-choice item, the panel members were asked to independently rate the percent of minimally competent beginning teachers whom they think who would likely answer the item correctly. For each constructed-response item, members were asked to independently indicate the level of response that would likely be achieved by the minimally competent beginning teacher.

- Round Two: The Round One ratings, which were displayed anonymously, were distributed, and members discussed the reasoning used in making their determinations. The second round moved the panel from individual item ratings to ratings at the section level (i.e., multiple-choice section and constructed-response section). They were asked the number of multiple-choice items that would be answered correctly and the total score points that would likely be achieved on the constructed-response items by the minimally competent beginning teacher.

- Round Three: Panel members were given the results of their Round Two ratings, along with information about the examinee pass rates at various panel member ratings. They

were then asked to make final independent recommendations for a passing standard based on the raw score points earned on each section of the test.

Separate ratings for each of the three subtests were made during each of the three rounds. The Panel's recommendation represents the computed median of the third round results.

Results of the Standard Setting Study
Because of the timing of the deadlines for Commission agenda items relative to the timing of the standard setting study on February 3, 2015, it was not possible to present the standard setting results in this agenda item. The results along with a staff recommendation will be provided in an in-folder item.

Standard Error of Measurement (SEM)
Once the final panel score recommendation is determined, an additional modification may be made to that score before it is recommended to the Commission. This modification is the determination and application of an adjustment that takes into consideration the Standard Error of Measurement (SEM). The SEM is a key measurement concept that addresses how accurately the recommended passing score standard reflects the scores likely to be achieved by actual candidates in real-world testing situations. For example, an examinee takes the test one time and receives a score. But if that same examinee were to take the same exam several times, with no change in his or her level of knowledge and preparation, it is possible that some of the resulting scores would be slightly higher or slightly lower than the score initially achieved by the examinee the first time he or she took the examination. Given this variation in possible scores on the same test by the same examinee, the examinee's initial score might not reflect the best score that examinee would hypothetically be able to achieve based on his or her actual knowledge and ability in the content area.

The range of scores an examinee would achieve across multiple administrations of the same test, were this activity to take place, includes what is known as the examinee's "true" score (i.e., the hypothetical score that would best reflect the examinee's actual ability) and the "observed score" (i.e., the actual score received on the first test administration).

A simple way to look at the concept of the standard error of measurement is to consider the case of the examinee who takes the CSET: Multiple Subjects examination one time. Many factors affect how the examinee does on his/her first attempt on the test, including knowledge of the content tested, affective factors such as the examinee's emotional, physical, and/or mental state on that particular day and time, and external factors such as the testing environment. Thus, it is not possible to say with certainty that the score obtained on the initial test taken by the examinee most accurately reflects his/her true level of knowledge, skills, and abilities. The likelihood that the examinee's true score is reflected on his/her first attempt is unknown. Thus, a computed Standard Error of Measurement is typically applied to adjust the minimum passing score for an examination in order to account for the difference in the examinee's true score and the examinee's observed score on the assessment.

How Does Applying the SEM Work?

As noted above, individual examinee scores on the first attempt could potentially not represent the examinee's true level of knowledge, skills, and abilities. The score could also potentially represent a "false negative" (i.e., the examinee did have sufficient knowledge, skills, and abilities but the actual score did not closely enough represent the examinee's true abilities) or a "false positive" (i.e., the examinee did not actually have sufficient knowledge, skills, and abilities but was able to earn a higher score than otherwise warranted). In the case of false positives, there is no psychometric approach to mitigating this outcome; in other words, the examinee is fortunate in achieving the passing score. However, there is a standard psychometric technique that does address the case of false negatives. This technique is the application of the SEM to the passing score standard established for a particular examination.

For example, on the CSET examination series, an examinee earns a raw score that is then translated into a scaled score. The SEM would be applied to the raw score that equates to the Commission's adopted passing score standard of a scaled score of 220. Thus, if the recommended cut score for an exam were to be a raw score of 30, which would equate to the adopted scale score of 220, the SEM would be applied to the raw score of 30. If the calculated SEM was minus 2 raw score points, and was applied to the raw cut score of 30, the raw cut score would now be 28. Examinees would need to achieve at least 28 raw score points to pass the examination. If an examinee whose actual knowledge and ability should have allowed him or her to pass was only able to earn 29 raw score points due to factors other than his/her knowledge of the content such as, for example, emotional upset, application of the SEM to the minimum passing standard would allow him or her to receive a passing score, thereby avoiding a false negative.

The SEM can range depending on the nature of the particular examination and the range of the candidate population for that examination. Typically multiple choice examinations that have clear right or wrong responses will have less variability in the range of candidate scores – either the candidate knows or does not know the content being assessed. In the case of constructed response and performance items, where candidates construct their own responses which are scored by trained readers, one might expect a larger range of variability in both responses and the background knowledge and abilities of candidates. It might also differ in the case where an examination is new, or the number of examinees is very low. Thus, an SEM could range from -1 to -5 or even higher. Each SEM is calculated individually and applied to the passing score for each examination.

The in-folder item to this agenda item will contain the recommended passing score standard for the CSET: Mathematics examinations and a recommendation concerning the application of an SEM to the raw score which equates to the scaled score of 220.

Next Steps

If the Commission adopts the recommended passing standards for the CSET: Mathematics examinations, notification will be posted on the CSET website and distributed to the field as soon as possible. In addition, recent examinees' scores will be tabulated based on the adopted

passing standard and scaled to a range of 100 to 300, with 220 representing the adopted passing standard for each exam. Individual examinee score reports will then be distributed within three to four weeks of the Commission's decision. The passing standard adopted by the Commission will also be applied to all subsequent administrations.

Appendix A
Outline of the CSET: Mathematics Subject Matter Requirements

The complete CSET: Mathematics Subject Matter Requirements, including the extensive descriptive text for each of the competencies can be found on the CTC exams website or by following this link: http://www.ctcexams.nesinc.com/PDF/CSET_Prep/CS_mathematics_SMR.pdf.

Content Domains for Subject Matter Understanding and Skill in Mathematics

Domain 1: Number and Quantity
1.1 The Real and Complex Number Systems
1.2 Number Theory

Domain 2: Algebra
2.1 Algebraic Structures
2.2 Polynomial Equations and Inequalities
2.3 Functions
2.4 Linear Algebra

Domain 3: Geometry
3.1 Plane Euclidean Geometry
3.2 Coordinate Geometry
3.3 Three-Dimensional Geometry
3.4 Transformational Geometry

Domain 4: Probability and Statistics
4.1 Probability
4.2 Statistics

Domain 5: Calculus
5.1 Trigonometry
5.2 Limits and Continuity
5.3 Derivatives and Applications
5.4 Integrals and Applications
5.5 Sequences and Series

Appendix B
Performance Characteristics and Scoring Scales for Constructed Response Items

CSET: Mathematics Performance Characteristics and Three-Point Score Scale
(for "focused" constructed-response items)

Performance Characteristics

Performance Characteristic	Description
PURPOSE	The extent to which the examinee responds to the constructed-response assignment's charge in relation to relevant CSET subject matter requirements.
SUBJECT MATTER KNOWLEDGE	The application of accurate subject matter knowledge as described in the relevant CSET subject matter requirements.
SUPPORT	The appropriateness and quality of the supporting evidence in relation to relevant CSET subject matter requirements.
DEPTH AND BREADTH OF UNDERSTANDING	The degree to which the response demonstrates understanding of the relevant CSET subject matter requirements.

Score Scale

Score	Score Point Description
4	The "4" response reflects a thorough command of the relevant knowledge and skills as defined in the subject matter requirements for CSET: Mathematics. The purpose of the assignment is fully achieved. There is a substantial and accurate application of relevant subject matter knowledge. The supporting evidence is sound; there are high-quality, relevant examples. The response reflects a comprehensive understanding of the assignment.
3	The "3" response reflects a general command of the relevant knowledge and skills as defined in the subject matter requirements for CSET: Mathematics.. The purpose of the assignment is largely achieved. There is a largely accurate application of relevant subject matter knowledge. The supporting evidence is adequate; there are some acceptable, relevant examples. The response reflects an adequate understanding of the assignment.

2	The "2" response reflects a limited command of the relevant knowledge and skills as defined in the subject matter requirements for CSET: Mathematics. The purpose of the assignment is partially achieved. There is limited accurate application of relevant subject matter knowledge. The supporting evidence is limited; there are few relevant examples. The response reflects a limited understanding of the assignment.
1	The "1" response reflects little or no command of the relevant knowledge and skills as defined in the subject matter requirements for CSET: Mathematics. The purpose of the assignment is not achieved. There is little or no application of relevant subject matter knowledge. The supporting evidence is weak; there are no or few relevant examples. The response reflects little or no understanding of the assignment.
U	The "U" (Unscorable) is assigned to a response that is unrelated to the assignment, illegible, primarily in a language other than English, or does not contain a sufficient amount of original work to score.
B	The "B" (Blank) is assigned to a response that is blank.