

*California  
Commission on Teacher Credentialing*

*Meeting of  
November 5-6, 2003*

**AGENDA ITEM NUMBER:**           **PERF - 2**

**COMMITTEE:**                       **Performance Standards**

**TITLE:**                               **Recommended Passing Score for the California  
Teaching Performance Assessment (CA TPA)**

**Action**

**Information**

**Report**

**Strategic Plan Goal(s):**

- Goal 1:       Promote educational excellence through the preparation and certification  
of professional educators**
- Sustain high quality standards for the preparation of professional educators
  - Sustain high quality standards for the performance of credential candidates

**Presented By:**       **Amy Jackson**

**Prepared By:**       \_\_\_\_\_                       **Date:** \_\_\_\_\_  
**Marjorie Suckow**  
**Consultant, Professional Services Division**

**Prepared By:**       \_\_\_\_\_                       **Date:** \_\_\_\_\_  
**Amy Jackson**  
**Administrator, Professional Services Division**

**Approved By:**       \_\_\_\_\_                       **Date:** \_\_\_\_\_  
**Beth Graybill**  
**Interim Director, Professional Services Division**

**Authorized By:**       \_\_\_\_\_                       **Date:** \_\_\_\_\_  
**Dr. Sam W. Swofford**  
**Executive Director**



# **Recommended Passing Score for the California Teaching Performance Assessment (CA TPA)**

## **Professional Services Division**

**November 5-6, 2003**

### **Executive Summary**

SB 2042 (Alpert/Mazzoni, Chapter 548, Statutes of 1998) requires all Preliminary Teaching Credential candidates to pass a teaching performance assessment (TPA). Professional teacher preparation programs may use the California Teaching Performance Assessment (CA TPA) developed by the Commission or may develop their own performance assessment. All teaching performance assessments are designed based on standards described in the Standards of Quality and Effectiveness for Professional Teacher Preparation Programs and should measure the Teaching Performance Expectations (TPEs) listed in the Professional Preparation Program Standards.

In developing CA TPA, the Commission contracted with the Educational Testing Service (ETS) in 2002-03 to develop four performance assessment tasks, a candidate handbook, task-specific rubrics, a record of evidence (ROE), benchmark cases, independent scoring cases, assessor training materials and information resources for the institutions that are implementing the CA TPA. ETS conducted a standard-setting study in June 2003 to provide the Commission with recommendations, based on the informed judgments of California educators and researchers, relevant to the determination of a passing score for the CA TPA.

This report provides background information on the development of the CA TPA, describes the CA TPA standard-setting study and the results of that study, and includes the staff recommendation for action.

### **Fiscal Impact Summary**

Title II funds supported the development the CA TPA system and the standard-setting study described in this report.

### **Policy Issues to be Considered**

What passing score should be recommended to the institutions that implement the CA TPA?

### **Recommendation**

Staff recommends that the Commission adopt a passing score for the CA TPA. The overall recommended CA TPA passing score is 12.



# **Recommended Passing Score for the California Teaching Performance Assessment (CA TPA)**

## **Professional Services Division**

**November 5-6, 2003**

### **Part I: Background information on the development of California Teaching Performance Assessment (CA TPA)**

California Senate Bill 2042 (Alpert/Mazzoni, Chapter 548, Statutes of 1998) requires all Preliminary Teaching Credential candidates to pass a teaching performance assessment (TPA). Professional teacher preparation programs may use the California Teaching Performance Assessment (CA TPA) developed by the California Commission on Teacher Credentialing (CCTC) or preparation programs may develop their own performance assessment. All teaching performance assessments are designed based on standards described in the Standards of Quality and Effectiveness for Professional Teacher Preparation Programs and should measure the Teaching Performance Expectations (TPEs) listed in the Professional Preparation Program Standards.

Educational Testing Service (ETS) was contracted from June 1, 2002 to August 31, 2003 to develop four performance assessment tasks, a candidate handbook, task-specific rubrics, a record of evidence (ROE), benchmark cases, independent scoring cases, assessor training materials and information resources for the institutions that are implementing the CA TPA.

#### **CA TPA tasks**

The CA TPA system consists of four separate tasks, with each focused on a different aspect of teaching practice. These tasks are inter-related yet separate and increase in complexity as the teacher candidates move through them. Collectively, these tasks measure aspects of the 13 TPEs that describe what all California beginning teachers need to know or be able to do to qualify for Preliminary Multiple Subject and Single Subject Credentials. The four tasks are as follows:

**Task 1: Principles of Content-Specific and Developmentally Appropriate Pedagogy** - teacher candidates are given the opportunity to demonstrate their knowledge of principles of developmentally appropriate pedagogy, of specific pedagogical skills for subject matter instruction as well as interpretation and use of assessment, and of adaptation of content for English learners and students with special needs.

**Task 2: Connecting Instructional Planning to Student Characteristics for Academic Learning** - teacher candidates are given the opportunity to demonstrate their ability to learn important details about a small group of learners and to design a lesson that is shaped by those contextual details.

**Task 3: Classroom Assessment of Academic Learning Goals** - teacher candidates are given the opportunity to demonstrate their ability to design standards-based, developmentally appropriate student assessment activities in the context of a small group of students and a specific lesson. In addition, candidates demonstrate their ability to assess student learning and to diagnose student needs from individual responses to the assessment activities.

**Task 4: Academic Lesson Design, Implementation, and Reflection after Instruction** - teacher candidates are given the opportunity to demonstrate their ability to design a standards-based lesson for a class of students, implement that lesson making appropriate use of class time and instructional resources, meet the differing needs of individuals within the class, manage instruction and student interaction, assess student learning, and analyze the strengths and weaknesses of the lesson.

The development of the CA TPA system includes several phases – development of the scoring rubrics for each of the four tasks, pilot testing and field review of the tasks, training of assessors to uniformly apply the scoring criteria, and establishing a passing standard.

### **Task-specific rubrics**

There is one unique scoring rubric for each task. There is a maximum of four points for each task, with 4.0 being the highest score level for a candidate's performance and 1.0 the lowest. The teacher candidate's response to each task will be judged on the relevance, accuracy, and appropriateness of evidence submitted. The four levels are outlined below in descending order:

**Level 4** - The response provides evidence that clearly, consistently, and convincingly demonstrates the teacher candidate's ability to... The preponderance of evidence provided for each of the following domains is appropriate, relevant, accurate, and clear or detailed. Evidence is purposefully connected and reinforced across the response.

**Level 3** - The response provides evidence that clearly demonstrates the teacher candidate's ability to... The preponderance of evidence provided for each of the following domains is appropriate, relevant, or accurate. Evidence is connected across the response.

**Level 2** - The response provides evidence that partially demonstrates the teacher candidate's ability to... The preponderance of evidence provided for each of the following domains is minimal, limited, cursory, inconsistent, and/or ambiguous. Evidence is weakly connected across the response and may be inconsistent.

**Level 1** - The response provides evidence that does little or nothing to demonstrates the teacher candidate's ability to... The preponderance of evidence provided for each of the following domains is inappropriate, irrelevant, inaccurate, or missing. Evidence is unconnected across the response.

## **Focus review groups**

The Commission and ETS established two focus review groups to assist with the development and tryout of the CA TPA as well as the TPE scales, task-specific rubrics, candidate handbook, and feedback forms. Based on the feedback from the focus groups, the Commission and ETS modified the tasks. All items went through the ETS sensitivity and fairness review process before being piloted.

## **Pilot test**

The Commission and ETS conducted a pilot test of the CA TPA from February to May 2002. The purpose of the pilot test was to collect information about the tasks, reactions to the tasks, and recommendations for modifying the tasks. Each of the four tasks was separately pilot-tested by different groups of participants. The Commission and ETS invited a few other programs to join members of the focus review groups to assist with the formative scoring sessions of the CA TPA Pilot Test. There were two formative scoring sessions held in April/May 2002 in Oakland. Based on the input from the participants, all four tasks, scales, and the candidate handbook were revised before the field review. Because 13 scales, one for each TPE, were found to be too complex to use for scoring, task-specific rubrics were developed for the field review.

## **Field review**

The Commission and ETS conducted a field review of the CA TPA from October 2002 through April 2003 via the Commission's website. Ninety-four institutions of higher education (IHEs) were invited to have their teacher candidates participate in the field review of the CA TPA. A total of 516 candidates from 19 IHEs downloaded materials from the website. Among this group, a total of 402 candidates from 18 IHEs submitted responses to one or more tasks, and 200 from 16 IHEs submitted responses to all four tasks. The characteristics of the teacher candidates and list of IHEs by type of programs are presented in Appendix A and B, respectively.

## **Scoring of task responses from the field review**

Only the responses from teacher candidates who completed all four tasks were scored at the field review scoring session. A total of 104 responses were scored across the four tasks. Each task was scored on a 1.0-to-4.0 scale, based on the task-specific rubrics. The distribution of scores varied somewhat across the tasks. For example, Task 2 had the largest proportion of candidates scoring a 1.0 (14%), while Task 4 had the smallest (3%). Task 3 had the largest proportion scoring a 4.0 (11%). For all four tasks, the highest frequency distribution was observed for score level 2.0: 31%, 25%, 22%, 30%, respectively. The frequency distribution of scores on each task is presented in Appendix C.

## **Benchmark and independent scoring cases**

In addition to the scoring of task responses, the field review activities included the selection of Benchmark and Independent Scoring Cases from candidate responses, development of guidelines for implementation of the CA TPA system, and drafting of training materials for TPA administrators and assessors. Establishing benchmarks for each score level of each task was an important step in establishing validity for the CA TPA.

From April 28 to May 2, 2003, ETS and the Commission conducted a work session to select benchmark and independent scoring cases from a representative sampling of teacher candidate responses. Twenty-six members from 22 IHEs completed this critical work. For each task, participants compiled a set of annotated benchmark cases with record of evidence forms as well as a set of independent scoring cases, also with record of evidence forms. Participants in Benchmarking became very knowledgeable about all aspects of assessor training and received all training materials. Several of those participants then functioned as co-trainers in the statewide Training of Lead Assessors/Scoring/Standard-setting session that was held on June 9-14, 2003 in Concord, California.

## **Lead assessor training**

The last phase of the development for the CA TPA system included the training of individuals to serve as lead assessors for their teacher preparation programs candidates, the scoring of teacher candidates' completed field review responses to the tasks, and participate in a standard setting study. Participants not only assisted in establishing the validity and reliability of the CA TPA system, but also benefited from gaining a comprehensive understanding of the CA TPA, the assessment scoring system, and how to train individuals who will eventually serve as assessors within a program.

During this six-day workshop, participants received foundation training in the TPEs, history and development of the CA TPA, assessment principles, the tasks, the task-specific rubrics, bias awareness and evidence-based scoring. Participants also received training for, and an opportunity to practice, scoring a task using benchmark and independent scoring cases. In addition, they participated in a standard-setting study for the CA TPA.

## Part II: Standard-setting Study for the CA TPA

The purpose of the standard-setting study for the CA TPA was to gather recommendations for the minimum, total CA TPA score, believed necessary for teacher candidates to meet, in order to pass the CA TPA. A total of 43 assessors, who participated in the CA TPA field review scoring session, took part in the standard-setting study for the CA TPA. Appendix D provides the demographics of the assessors (Table 4) and type of institutions they represented (Table 5).

The standard-setting study was conducted during the last two days of the scoring session. As part of the assessor orientation session, assessors discussed the concept of minimal competence. Candidates who demonstrated “just enough” TPE-related knowledge and skills to perform important teaching tasks satisfactorily, as applied to the CA TPA, defined minimal competence. Characteristics of minimal competence were noted on chart paper, a summary shared with the assessors and then modified based on their comments. The revised definition served as the frame-of-reference for the subsequent standard-setting judgments and discussions.

On the fifth day of the scoring session, two ETS research staff conducted the first part of the standard-setting process in each of the task-specific scoring rooms. The Assessors’ task was described as making a recommendation of the minimum score teacher candidates need to earn on the CA TPA, in order to receive a preliminary credential to practice. In other words, they were asked to decide how much TPE-related knowledge or skill was absolutely necessary to signify that a candidate was ready to enter the teaching profession. The Commission had made a policy decision that a teacher candidate could not pass the CA TPA with a score of 1.0 on any task, regardless of scores on the other tasks.

Each assessor was asked to indicate, independently, the level of task-specific performance (rubric score value) that would most likely be achieved by a minimally competent teacher candidate. Each cadre of assessors went through two iterations of judgments. The first round was made independently once the training was completed, and the second was made after assessors had an opportunity to share their rationales for their first-round judgments. These task-level averages were summed to provide a preliminary overall TPA passing score. The mean and median selected scores by tasks are summarized in Appendix E. *(Please refer to Table 6 and 7: Results from First and Second Rounds of Judgments from the Standard-setting Study.)*

On the sixth day, the assessors assembled together and spent approximately three hours presenting (in task-specific groups) their task and understanding of the rubric to the other task groups. After the review of the four CA TPA tasks, the assessors were presented with a summary of the standard-setting results of the previous day. The assessors were asked to discuss their reactions to the initial overall CA TPA passing score recommendation, after which each assessor was asked to record his or her final overall recommended CA TPA passing score. The mean of the final recommended passing scores was 12 and the standard error of measurement was 0.79. *(Please refer to Appendix E – Table 8 for Distribution of Passing Scores.)*

### **Part III: Recommendation for Overall CA TPA Passing Score**

The mean of the final distribution of passing scores, 12 may be taken as the overall recommended CA TPA passing score. The lowest possible score on the CA TPA is 8 points (as a task-level score of 1.0 is not acceptable) and the highest possible score is 16 points. The Commission is encouraged to consider the mean summary value of the recommended CA TPA passing score, the standard error of measurement, and the standard-setting process used to arrive at this value as it establishes the operational CA TPA passing score.

**Appendix A**  
**Table 1: Teacher Candidate Demographics**

|                                  | Candidates who submitted responses to all four tasks |         |
|----------------------------------|--|---------|
|                                  | Number   | Percent |
| <b>Gender</b>                    |  |         |
| Female                           | 151  | 75.5%   |
| Male                             | 49   | 24.5%   |
| <b>Race or ethnicity</b>         |  |         |
| African American                 | 6  | 3.0%    |
| American Indian/Alaskan Native   | 4  | 2.0%    |
| Filipino                         | 5  | 2.5%    |
| Hispanic                         | 30   | 15.0%   |
| Pacific Islander                 | 1  | 0.5%    |
| White                            | 126  | 63.0%   |
| Other                            | 28   | 14.0%   |
| <b>Type of teacher candidate</b> |  |         |
| Multiple Subject                 | 146  | 73.0%   |
| Single Subject                   | 54   | 27.0%   |
| <b>School level</b>              |  |         |
| Elementary                       | 137  | 68.5%   |
| Middle                           | 21   | 10.5%   |
| High                             | 42   | 21.0%   |
| <b>Subject area</b>              |  |         |
| English/Language Arts            | 18   | 9.0%    |
| Mathematics                      | 15   | 7.5%    |
| Science                          | 12   | 6.0%    |
| History/Social Science           | 9  | 4.5%    |
| Multiple Subjects                | 146  | 73.0%   |



## Appendix B

**Table 2: Participating Programs (with candidates submitting all four tasks)**

| Institution of Higher Education (IHEs) | Type of Program                     | Number of<br>teacher candidates<br>submitting all<br>four tasks |
|--|-------------------------------------|---|
| California Lutheran University         | Post-baccalaureate                  | 6   |
| CalStateTEACH                          | Internship                          | 22  |
| Chapman University                     | Post-baccalaureate                  | 2   |
| CSU, Bakersfield                       | Post-baccalaureate, Blended, Intern | 19  |
| CSU, Hayward                           | Post-baccalaureate, Blended, Intern | 17  |
| CSU, Long Beach                        | Post-baccalaureate, Blended, Intern | 15  |
| CSU, Pomona                            | Post-baccalaureate, Intern          | 3   |
| CSU, San Marcos                        | Post-baccalaureate, Blended, Intern | 2   |
| LAUSD                                  | Internship                          | 29  |
| Loyola Marymount University            | Post-baccalaureate, Intern          | 5   |
| Ontario-Montclair District Program     | Internship                          | 1   |
| Pepperdine University                  | Post-baccalaureate                  | 3   |
| San Joaquin COE                        | Internship                          | 28  |
| UC, Riverside                          | Post-baccalaureate, Blended, Intern | 13  |
| UC, Santa Barbara                      | Post-baccalaureate                  | 6   |
| Vanguard University                    | Post-baccalaureate                  | 29  |



## Appendix C

**Table 3: Frequency Distribution of Scores on each Task**

| Score* | Task 1 | Task 2 | Task 3 | Task 4 |
|--------|--------|--------|--------|--------|
| 1.0    | 8%     | 14%    | 7%     | 3%     |
| 1.5    | 6%     | 17%    | 13%    | 13%    |
| 2.0    | 31%    | 25%    | 22%    | 30%    |
| 2.5    | 26%    | 18%    | 18%    | 23%    |
| 3.0    | 15%    | 18%    | 19%    | 16%    |
| 3.5    | 10%    | 4%     | 10%    | 7%     |
| 4.0    | 5%     | 3%     | 11%    | 8%     |

\*A candidate's score was calculated as the average of two assessors' scores, so half-point values were possible.  
*Note:* Percentages do not always sum to 100% due to rounding.



**Appendix D**  
**Table 4: Assessor Demographics**

|                                  | Number of<br>Assessors | Percent of<br>Assessors |
|----------------------------------|------------------------|-------------------------|
| <b>Gender</b>                    |                        |                         |
| Female                           | 29                     | 67%                     |
| Male                             | 14                     | 33%                     |
| <b>Ethnicity*</b>                |                        |                         |
| African American                 | 1                      | 2%                      |
| American Indian or Alaska Native | 1                      | 2%                      |
| Asian American                   | 2                      | 5%                      |
| Hispanic                         | 3                      | 7%                      |
| Other                            | 2                      | 5%                      |
| White                            | 32                     | 78%                     |
| <b>Current Occupation</b>        |                        |                         |
| Teacher Educator                 | 36                     | 84%                     |
| Administrator                    | 4                      | 9%                      |
| Advisor                          | 1                      | 2%                      |
| Other                            | 2                      | 5%                      |

*Note:* Percentages do not always sum to 100% due to rounding.

\*Two assessors did not indicate their ethnicity.

**Table 5: Type of Institution**

| <b>Type of Institution</b>  | Number of<br>Assessors | Percent of<br>Assessors |
|-----------------------------|------------------------|-------------------------|
| California State University | 14                     | 33%                     |
| Private University          | 23                     | 53%                     |
| Intern Program              | 5                      | 12%                     |
| Other                       | 1                      | 2%                      |



## Appendix E

**Table 6: Results from First Round of Judgments from the Standard-setting Study**

| Task   | Number of Assessors | Mean selected score | Median selected score | Standard deviation |
|--------|---------------------|---------------------|-----------------------|--------------------|
| Task 1 | 9                   | 2.7                 | 3.0                   | 0.5                |
| Task 2 | 9                   | 2.8                 | 3.0                   | 0.5                |
| Task 3 | 10                  | 2.9                 | 3.0                   | 0.3                |
| Task 4 | 15                  | 2.7                 | 3.0                   | 0.5                |

Combined total score (sum of task-level mean judgments) = 11.1

**Table 7: Results from Second Round of Judgments from the Standard-setting Study**

| Task   | Number of Assessors | Mean selected score | Median selected score | Standard deviation |
|--------|---------------------|---------------------|-----------------------|--------------------|
| Task 1 | 9                   | 3.0                 | 3.0                   | 0.0                |
| Task 2 | 9                   | 2.9                 | 3.0                   | 0.4                |
| Task 3 | 10                  | 3.0                 | 3.0                   | 0.0                |
| Task 4 | 15                  | 2.7                 | 3.0                   | 0.5                |

Combined total score (sum of task-level mean judgments) = 11.6

**Table 8: Distribution of Passing Scores**

| Recommended Passing Scores | Number of Assessors | Percentage of Assessors |
|----------------------------|---------------------|-------------------------|
| 8                          | 0                   | 0%                      |
| 9                          | 0                   | 0%                      |
| 10                         | 3                   | 7.5%                    |
| 11                         | 8                   | 20.0%                   |
| 12                         | 28                  | 70.0%                   |
| 13                         | 1                   | 2.5%                    |
| 14                         | 0                   | 0%                      |
| <hr/>                      |                     |                         |
| Mean (rounded)             | 12                  |                         |
| Median                     | 12                  |                         |
| Standard Deviation         | 0.66                |                         |

*Note:* The rounded mean of the assessors' ratings is used as the recommended passing score. Three assessors were not in attendance on Day 6.

