## Information

*Professional Services Committee*

# Continuation of the Discussion of the Teaching Performance Assessment

**Executive Summary:** This agenda item continues the discussion about the Teaching Performance Assessment and looks specifically at issues relating to the Commission's statutory management of the TPA process, including relevant standards, oversight processes, and the Commission's relationship with developers of approved TPA models.

**Policy Question**: This agenda item raises several policy issues for the Commission's discussion and potential future direction and/or action.

**Recommended Action:** For information only

**Presenters:** Phyllis Jacobson, Administrator, and Mike Taylor, Consultant, Professional Services Division

**Strategic Plan Goal**

*I. Educator Quality*
   ♦ Maintain expectations for educator preparedness and performance that are responsive to the needs of California's diverse student population and promote 21st century teaching and learning.
   ♦ 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.

September 2012

# Continuation of the Discussion of the Teaching Performance Assessment

**Introduction**
This agenda item continues the discussion of issues relating to the implementation of the Teaching Performance Assessment (TPA) requirement begun at the Commission's April 2012 meeting (http://www.ctc.ca.gov/commission/agendas/2012-04/2012-04-6B.pdf).

The first part of this item provides background information on the Commission's statutory role with respect to TPA implementation and management; Part II discusses potential policies pertinent to the design and implementation of the teaching performance assessment; and Part III discusses the Commission's relationship with developers of approved alternative teaching performance assessment models.

**Background**
As of July 2008, through SB 1209, California statute (Education Code §44320.2) requires all candidates for a Preliminary Multiple and Single Subject Teaching Credential to pass an assessment of their teaching performance with K-12 public school students as part of the requirements for earning a preliminary teaching credential. In addition to the mandate that the Commission develop, maintain, and provide training for its own Teaching Performance Assessment (TPA) model, the "CalTPA," the Commission has the authority to approve additional TPA models that meet the Commission's adopted Assessment Design Standards. To date, the Commission has approved two additional models, the Fresno Assessment of Student Teachers (FAST) and the Teaching Event component of the Performance Assessment for California Teachers (PACT).

Approximately 23,065 candidates took the TPA in 2010-2011. Of these, 63% took the CalTPA; 33.6% took the PACT; and the remaining 3.4% took the FAST.

**Part I: The Commission's Statutory Responsibilities With Respect to TPA Implementation and Oversight**
The Education Code (Appendix A) codifies an extensive range of responsibilities for the Commission, including  developing and implementing its own TPA model for use by any program in the state, establishing design standards and reviewing alternative TPA models against those standards, collecting and analyzing assessment data, assuring the examination is free of bias, and assuring that the TPA is embedded in programs, provides formative data, and serves as one requirement for recommending candidates for a credential.

To implement the Commission's responsibility for establishing design standards and reviewing alternative TPA models against those standards, a set of "Assessment Design Standards" (Appendix B) was developed and adopted initially by the Commission in 2001, prior to the implementation of SB1209 which both set the July 1, 2008 start date for the TPA requirement in California and envisioned the TPA as a program embedded, locally administered and scored

assessment. The two Assessment Design Standards were written and adopted at a time when it was expected that the TPA would be more of a standardized assessment that would be centrally administered and scored rather than a locally owned and scored assessment. The standards assumed that local preparation programs would be likely proposers of any alternative TPA rather than outside entities, which was a logical assumption at the time the standards were developed, since the TPA was a brand-new concept on the national scene and California was the pioneer in this mode and model of assessment required for all new credential candidates.

The Assessment Design Standards are based foundationally on the assessment quality standards represented by the *Joint Standards for Educational and Psychological Testing* of the American Educational Research Association, the American Psychological Association, and the National Council for Measurement in Education. These standards clearly outline requirements for assessment reliability, as well as for many other psychometric properties and requirements, that all Commission-approved TPA models should meet since the TPA is a large-scale state-mandated assessment. These standards are designed to assure that the properties of assessments that contribute to decisions about individual candidates are legally defensible. This aspect of the standards is also consistent with statutory requirements for valid and reliable TPA assessments.

### *A Closer Look at the Assessment Design Standards*
There are two assessment design standards adopted by the Commission. The first standard focuses on technical issues relating to the validity and fairness of the assessment design itself. The second standard focuses on technical issues relating to establishing and maintaining the fairness of the scoring process within local implementation of the assessment, in keeping with the assessment's design and technical characteristics.

For example, the first standard focuses primarily on:
- The Commission's requirement that the tasks of the TPA focus on the major domains of the Teaching Performance Expectations (TPEs)
- The development and periodic review of the validity and fairness of the assessment tasks and of the scoring scales
- Appropriate scoring of a variety of possible pedagogical practices across the different credential areas
- Scoring scales that focus primarily on teaching performance and minimize effects of other candidate factors not clearly related to pedagogical competence
- The requirement of a clear statement of the intended uses of the assessment, including appropriate cautions about additional or alternative uses for which the assessment is not valid
- Content review and editing procedures to assure that tasks and directions are clear to candidates from a variety of cultural and linguistic backgrounds
- The requirement for basic psychometric analyses to identify any potential differential effects in relation to candidates' race, ethnicity, language, gender, or disability; and the requirement to act promptly to maximize the fairness of the assessment for all groups of candidates in the event any bias is found

The second standard focuses primarily on issues related to local scoring of the assessment:
- The sufficiency of candidate evidence to yield an appropriate judgment of each candidate's pedagogical qualifications for a preliminary teaching credential
- A comprehensive and ongoing assessor training and retraining process to assure each scorer's accuracy
- Periodic evaluation of the assessor training process and its outcomes
- Rescoring of selected assessment tasks for the purpose of evaluating the reliability of scorers and the consistency of scoring over time
- Use of marker papers to facilitate training of first time and veteran scorers
- An appeal process for candidates unsuccessful on the assessment

The design standards have not been re-examined for some time. This is because implementation efforts to get the TPA off the ground in response to legislative requirements to begin implementation for all candidates in 2008 were monumental and all-consuming for both model developers and for local programs. It took an incredibly intensive effort to get a TPA system from zero to full implementation in a relatively short space of time, and to assure that all programs and all candidates were being appropriately served by and involved in the TPA process. Now that the TPA is fully implemented and programs are being supported in their implementation efforts, it time to assess where the Commission has been and where it needs/wants to go with respect to what the Assessment Design Standards should include.

### *Key Issues Relating to the Current Assessment Design Standards*
*A. Technical design requirements*
The current assessment design standards are still relevant and valid as far as their technical quality is concerned. They are still consistent with the psychometric principles embodied in the *Joint Standards for Educational and Psychological Testing* of the American Educational Research Association, the American Psychological Association, and the National Council for Measurement in Education, since these standards have not been modified since the several TPA models were developed and adopted.

It is less clear that there should be a distinction between standards one and two. In practice, the two standards are inextricably linked. The standards could be streamlined and better integrated into a single unified set.

*B. To Whom the Standards Apply*
The assessment design standards were written for what was assumed at the time to be a program sponsor who would individually submit an alternative assessment for review and possible implementation. In fact, the standards as written use the term "the sponsor of the professional teacher preparation program" rather than the "developer of the alternative teaching performance assessment model." However, over time, the sponsors of alternative TPA models have become not only a single preparation program sponsor (such as CSU Fresno, the developer of the FAST assessment), but also "virtual" sponsors such as the PACT Consortium, which consisted of Stanford University, all of the University of California campuses, plus some CSU and other postsecondary institutions. "Virtual" sponsors have no specific standing with the Commission, may not be incorporated specifically as a stand-alone entity, and do not participate in the Commission's accreditation system. There may well be additional future virtual sponsors of TPA

models. Given this situation, the assessment design standards might benefit from some revision to clarify who is eligible to submit a proposed TPA model for Commission review and approval, and what the characteristics of such entities must be.

*C. Who is responsible for the approved alternative models?*
Given that entities outside of the Commission's accreditation system submit TPA models that are and/or may be approved by the Commission in the future, and also the fact that some of these entities are not incorporated per se, it is not clear who is ultimately responsible for the ongoing technical support of the model once approved. In the case of the CalTPA, the Commission is ultimately responsible; in the case of FACT, CSU Fresno is ultimately responsible, but in the case of PACT it is less clear who is ultimately responsible. Although there is an entity known as "PACT Central," the standing of this entity and the Commission's relationship to this entity, which is outside of either an approved program or another element within the accreditation system, is not clearly defined.

*D. The Scope of the TPA Design Standards Relating to Multiple Subject Candidates*
An issue has been raised by some stakeholders regarding the multiple subject TPA in particular. The Assessment Design Standards require that each TPA model address all of the Commission-adopted TPEs. With respect to the Multiple Subject performance assessment, this requirement means that each candidate must be assessed in each of the four core content areas in relation to subject-specific pedagogy in order to address TPE 1, which addresses reading/language arts, mathematics, science and history/social science.

Some stakeholders have raised the question if it is necessary to require models to address all four of the core content areas for multiple subject candidates, while other stakeholders believe that since the TPEs address all four core content areas and the TPA is required to address all the TPEs, that the four core content areas should continue to be required for all multiple subject candidates. Also a factor in this issue is the upcoming implementation of the Common Core standards which focus heavily on English language arts and mathematics but not the other two core content areas. This situation has prompted some stakeholders to suggest that it might be appropriate under these circumstances to refocus the multiple subject TPA on literacy and mathematics core content areas for multiple subject candidates. However, there is also the question of whether, given the importance of subject-specific pedagogy in general and the fact that the Common Core standards address interrelated skills across the other content areas in addition to English and Math, it would be sufficient to address fewer than four core areas for multiple subject candidates within the TPA.

If the Commission were to choose to require fewer than four core content areas for existing as well as future TPA models for the multiple subject TPA, there could be design and cost implications to redesign the tasks and the scoring structure accordingly, and to retrain scorers for the expanded focus in these core content areas.

The policy issue for the Commission's discussion would be whether the TPA for multiple subject candidates should address all four core content areas, or fewer than the four.

*Given the discussion above, staff recommends:*

   1. *That the two assessment design standards be reviewed and potentially revised to focus exclusively on issues relating to the technical quality of the assessment instrument itself for purposes of initial review and approval by the Commission as a TPA model usable by California preliminary preparation programs.*
   2. *That the standards clarify, via Preconditions to be developed, who is eligible to sponsor an alternative TPA, and how that entity must interact with the Commission for oversight and reporting purposes.*
   3. *That the Commission determine if the design standards for Multiple Subject candidates should continue to require addressing all four core content areas, or fewer than four. If fewer than four, the Commission should specify which area(s) should be included.*

If the Commission were to revise the Assessment Design Standards, there would be implications for currently-approved models. For example:

   - Would currently-approved models need to be re-reviewed and approved?
   - Would currently-approved models need to redesign their tasks and other associated components if the Commission were to require fewer than four core content areas for Multiple Subject candidates? If so, would the models need to revalidate the revised tasks and scoring rubrics to assure both validity and fairness to candidates?

**Part II: Potential Additional Policies Relating to TPA Model Developers**
If the Commission were to decide to review and adopt revised TPA Assessment Design Standards as recommended above, the Commission might also want to consider whether any additional sets of policy requirements relating to the TPA would be necessary and/or appropriate.

For example, currently there is no formalized set of requirements for model developers once the models have been reviewed and approved against the two current Assessment Design Standards. Similarly, there is no formal reporting requirement or process when model developers make a modification to the model as it was originally reviewed and approved. Models do not necessarily remain static over time, as ongoing improvements are often necessary. For example, the CalTPA needed to develop additional single subject tasks for the Subject-Specific Pedagogy tasks for 13 content areas; PACT has also developed additional subject area-specific tasks. These ongoing modifications were not required to be reported in any formal way to the Commission. In general, model developers have communicated this information on an informal basis via discussions at the Users Advisory Committee (UAC) (see below for further information on the UAC and its activities to date).

Nor is there any formalized set of requirements for model developers' ongoing support of programs that are implementing one or more of the approved TPA models at the local level. Programs implementing the CalTPA have a vehicle through the CalTPA Steering Committee to indicate their support needs and to resolve implementation issues, and there is one dedicated PSD staff member who works with the CalTPA. FAST is a single entity that is both model sponsor and program implementer, so this issue is less of a concern for that model. PACT Central provides ongoing program support, but it is less clear what services are provided and how PACT Central interacts with local programs implementing the model.

Periodic updates from both model developers and local programs have been provided on an informal basis within agenda items addressing TPA updates, but these are not consistent in terms of the range and recency of information available to the Commission. The Commission might want to specify a more consistent reporting requirement and/or function from both model developers and local programs. Some of the reporting requirements could derive from activities in the current Assessment Design Standards that are more appropriate for the ongoing implementation of the model rather than characteristic of the psychometric properties of the assessment itself. For example:

- Acting promptly to identify any potential sources of bias and taking action to address that bias
- Periodically reconsidering the reasonableness of the scoring scales and the established passing standard
- Providing additional training to assessors if and when new pedagogical tasks and/or scoring scales are incorporated into the assessment

Any or all of these areas might benefit from the development of new policies, which might also address how these issues relate to the Commission's accreditation process for local programs. However, if the Commission were interested in moving forward with the development of any potential new standards relating to the TPA, it would also be important to determine:

- whether models would be required to document how they were meeting these standards on a periodic basis in order for the model to continue to be eligible for use in California
- how the Commission would monitor each model's continued compliance with a potential second set of standards
- the effects of the current fiscal crisis on model developers' abilities to implement additional standards
- how local programs would be required to document and/or report their implementation efforts, and what additional resources would be required for programs to undertake these activities beyond what they are already doing for accreditation purposes

*Considering the above factors, staff recommends that the Commission consider developing policies that address:*

1. *How model developers communicate with, support, and monitor local preparation programs implementing that model, and*
2. *How model developers gather implementation information from local program interactions and monitoring activities, and communicate that information to the Commission.*

If the Commission were to choose to develop and adopt any additional sets of TPA-related policies, some development work, including input from stakeholders would be needed. An appropriate process for such development could be to work with an advisory group of psychometric experts along with veteran TPA practitioners from each model.

**Part III: Structure for Commission Interaction with Model Developers**
Ongoing challenges, especially given the current budget crisis, include addressing how the Commission interacts with the model developers, how the model developers in turn interact with the local preparation programs which are required to embed the model into the fabric of their

preparation programs, and how the preparation programs interact with the model developers as well as with the Commission to provide the necessary information for the Commission to meet its statutory requirements.

In the early days of statewide implementation of the TPA, the Commission instituted a communication structure in 2008 that included representatives of not only model owners and program users, but also key stakeholders. This committee, the TPA Implementation Task Force, was set up to deal with immediate implementation questions that transcended individual models, such as strategies for implementation of the TPA under the budget crisis, samples of program uses of TPA results for induction purposes, communications with stakeholders about the TPA requirement, and similar topics. This group made recommendations to the Commission, one of which resulted in the development of several TPA public information leaflets still available on the Commission's website (http://www.ctc.ca.gov/educator-prep/TPA.html). This group met several times during 2009 and eventually formed the basis of what was to become the TPA Users Advisory Committee (UAC).

The UAC is comprised of model developers, the model's psychometricians, and program user representatives. Since 2010, the UAC met regularly to discuss issues related to implementation, data collection, and other topics which affect each of the TPA models. Due to the Commission's current budget situation, meetings of the UAC have currently been suspended. Thus, at the present time, the Commission does not have a formal vehicle for communication with and among the three approved model developers, psychometricians, and program users.

Outside of the context of discussions with the UAC, the Commission has not formally asked model developers specifically about how they are carrying out their support relationships with programs implementing the model, how they are modifying and/or adapting the model over time from the form initially approved by the Commission, and how they are responding to requests from program users for any variance in implementation from the model as originally approved.

In 2010, Commission staff began collecting annual TPA passing status data for all candidates from all programs across all models. Although the data collection process requires communication from the Commission to programs, it bypasses the model developers in the process, as individual programs upload data directly to the Commission for review, analysis, and reporting.

The model developers themselves communicate with program users in different ways. For example, the CalTPA established a Steering Committee with representatives of CalTPA program user representatives. The CalTPA Steering Committee has discussed and/or addressed the following types of items in collaboration with Commission staff:
- need for development of subject specific pedagogy tasks each of the subject content areas;
- training assessors throughout the state;
- implementation of an online calibration system for assessors;
- development of an online CalTPA resource center for TPA coordinators;
- rewriting the initial candidate handbook and implantation manual;

- updating benchmark and independent score cases to be used for initial assessor calibration;
- annual CalTPA coordinator conferences.

PACT Central operates as a central point of contact for the PACT model, providing guidance to program users and holding an annual users conference. Since FACT is a self-contained single user model, similar activities are conducted by CSU Fresno for its own model.

*Staff suggests that there needs to be an ongoing relationship between the model developers and the Commission. To this end, staff further suggests that the Commission formalize and provide continued support to the Users Advisory Committee and that the Commission clarify its expectations for the UAC's purpose and responsibilities. One of the UAC's next tasks could be to gather input from the model developers and representative model users regarding the review and revision of the Assessment Design Standards for future use, and to provide input into potential new policies regarding model developers' responsibilities with respect to local programs as well as to the Commission.*

**Next Steps and Future Agenda Items**
Based on Commission discussion and direction, staff will develop and present future agenda items related to the teaching performance assessment and/or the assessment design standards for Commission review and potential action.

# Appendix A

## Education Code Relating to the Teaching Performance Assessment

California Education code section 44320.2 - TPA

(a) The Legislature finds and declares that the competence and performance of teachers are among the most important factors in influencing the quality and effectiveness of education in elementary and secondary schools.

(b) Commencing July 1, 2008, for a program of professional preparation to satisfy the requirements of paragraph (3) of subdivision (b) of Section 44259, the program shall include a teaching performance assessment that is aligned with the California Standards for the Teaching Profession and that is congruent with state content and performance standards for pupils adopted pursuant to subdivision (a) of Section 60605. In implementing this requirement, institutions or agencies may do the following:

　(1) Voluntarily develop an assessment for approval by the commission. Approval of any locally developed performance assessment shall be based on assessment quality standards adopted by the commission, which shall encourage the use of alternative assessment methods including portfolios of teaching artifacts and practices.

　(2) Participate in an assessment training program for assessors and implement the commission developed assessment.

(c) The commission shall implement the performance assessment in a manner that does not increase the number of assessments required for teacher credential candidates prepared in this state. Each candidate shall be assessed during the normal term or duration of the preparation program of the candidate.

(d) Subject to the availability of funds in the annual Budget Act, the commission shall perform all of the following duties with respect to the performance assessment:

　(1) Assemble and convene an expert panel to advise the commission about performance standards and developmental scales for teaching credential candidates and the design, content, administration, and scoring of the assessment. At least one-third of the panel members shall be classroom teachers in California public schools.

　(2) Design, develop, and implement assessment standards and an institutional assessor training program for the sponsors of professional preparation programs to use if they choose to use the commission developed assessment.

　(3) Establish a review panel to examine each assessment developed by an institution or agency in relation to the standards set by the commission and advise the commission regarding approval of each assessment system.

　(4) Initially and periodically analyze the validity of assessment content and the reliability of assessment scores that are established pursuant to this section.

　(5) Establish and implement appropriate standards for satisfactory performance in assessments that are established pursuant to this section. The commission shall ensure that oral proficiency in English is a criterion for scoring the performance of each candidate in each assessment.

　(6) Analyze possible sources of bias in the performance assessment and act promptly to eliminate any bias that is discovered.

(7) Collect and analyze background information provided by candidates who participate in the performance assessment, and report and interpret the individual and aggregated results of the assessment.

(8) Examine and revise, as necessary, the institutional accreditation system pursuant to Article 10 (commencing with Section 44370) , for the purpose of providing a strong assurance to teaching candidates that ongoing opportunities are available in each credential preparation program that is offered pursuant to Section 44320, Article 6 (commencing with Section 44310) , Article 7.5 (commencing with Section 44325) , or Article 3 (commencing with Section 44450) of Chapter 3 for candidates to acquire the knowledge, skills, and abilities measured by the assessment system.

(9) Ensure that the aggregated results of the assessment for groups of candidates who have completed a credential program are used as one source of information about the quality and effectiveness of that program.

(e) The commission shall ensure that each performance assessment pursuant to subdivision (b) is state approved and aligned with the California Standards for the Teaching Profession and is consistently applied to candidates in similar preparation programs. To the maximum feasible extent, each performance assessment shall be ongoing and blended into the preparation program, and shall produce the following benefits for credential candidates, sponsors of preparation programs, and local education agencies that employ program graduates:

(1) The performance assessment shall be designed to provide formative assessment information during the preparation program for use by the candidate, instructors, and supervisors for the purpose of improving the teaching knowledge, skill, and ability of the candidate.

(2) The performance assessment results shall be reported so that they may serve as one basis for a recommendation by the program sponsor that the commission award a teaching credential to a candidate who has successfully met the performance assessment standards.

(3) The formative assessment information pursuant to paragraph (1) and the performance assessment results pursuant to paragraph (2) shall be reported so that they may serve as one basis for the individual induction plan of the new teacher pursuant to Section 44279.2.

(f) It is the intent of the Legislature that assessments in accordance with paragraphs (1) and (2) of subdivision (b), including the administrative costs of the commission, be fully funded.

**Amended by**: Stats. of 2005, Chapter 517 - §24   **Effective**: January 1, 2007
**Added by**: Stats. of 1997, Chapter 548 - §9   **Effective**: January 1, 1999

# Appendix B

## The Commission's Assessment Design Standards

**Assessment Design Standard 1: Assessment Designed for Validity and Fairness**
(Assessment Design Standard 1 Applies to Programs that Request Approval of Alternative Assessments)

The sponsor of the professional teacher preparation program requests approval of a Teaching Performance Assessment (TPA) in which complex pedagogical assessment tasks and multi-level scoring scales are linked to the Teaching Performance Expectations (TPEs). The program sponsor clearly states the intended uses of the assessment, anticipates its potential misuses, and ensures that local uses are consistent with the statement of intent. The sponsor maximizes the fairness of assessment design for all groups of candidates in the program, and ensures that the established passing standard on the TPA is equivalent to or more rigorous than the recommended state passing standard.

**Required Elements for Assessment Design Standard 1: Assessment Designed for Validity and Fairness**

1(a) The Teaching Performance Assessment includes complex pedagogical assessment tasks to prompt aspects of candidate performance that measure the TPEs. Each task is substantively related to two or more major domains of the TPEs. For use in judging candidate-generated responses to each pedagogical task, the assessment also includes multi-level scoring scales that are clearly related to the same TPEs that the task measures. Each task and its associated scales measure two or more TPEs. Collectively, the tasks and scales in the assessment address key aspects of the six major domains of the TPEs. The sponsor of the professional teacher preparation program documents the relationships between TPEs, tasks and scales.

1(b) To preserve the validity and fairness of the assessment over time, the sponsor may need to develop and field-test new pedagogical assessment tasks and multi-level scoring scales to replace or strengthen prior ones. Initially and periodically, the sponsor analyzes the assessment tasks and scoring scales to ensure that they yield important evidence that represents candidate knowledge and skill related to the TPEs, and serves as a basis for determining entry-level pedagogical competence to teach the curriculum and student population of California's K-12 public schools. The sponsor records the basis and results of each analysis, and modifies the tasks and scales as needed.

1(c) Consistent with the language of the TPEs, the sponsor defines scoring scales so different candidates for credentials can earn acceptable scores on the Teaching Performance Assessment with the use of different pedagogical practices that support implementation of the K-12 content standards and curriculum frameworks. The sponsor takes steps to plan and anticipate the appropriate scoring of candidates who use pedagogical practices that are educationally effective but not explicitly anticipated in the scoring scales.

1(d) The sponsor develops scoring scales and assessor training procedures that focus primarily on teaching performance and that minimize the effects of candidate factors that are not clearly related to pedagogical competence, which may include (depending on the

circumstances) factors such as personal attire, appearance, demeanor, speech patterns and accents that are not likely to affect student learning.

1(e) The sponsor publishes a clear statement of the intended uses of the assessment. The statement demonstrates the sponsor's clear understanding of the high-stakes implications of the assessment for candidates, the public schools, and K-12 students. The statement includes appropriate cautions about additional or alternative uses for which the assessment is not valid. Before releasing information about the assessment design to another organization, the sponsor informs the organization that the assessment is valid only for determining the pedagogical competence of candidates for initial teaching credentials in California. All elements of assessment design and development are consistent with the intended use of the assessment for determining the pedagogical competence of candidates for Preliminary Teaching Credentials in California.

1(f) The sponsor completes content review and editing procedures to ensure that pedagogical assessment tasks and directions to candidates are culturally and linguistically sensitive, fair and appropriate for candidates from diverse backgrounds. The sponsor ensures that groups of candidates interpret the pedagogical tasks and the assessment directions as intended by the designers, and that assessment results are consistently reliable for each major group of candidates.

1(g) The sponsor completes basic psychometric analyses to identify pedagogical assessment tasks and/or scoring scales that show differential effects in relation to candidates' race, ethnicity, language, gender or disability. When group pass-rate differences are found, the sponsor investigates to determine whether the differences are attributable to (a) inadequate representation of the TPEs in the pedagogical tasks and/or scoring scales, or (b) overrepresentation of irrelevant skills, knowledge or abilities in the tasks/scales. The sponsor acts promptly to maximize the fairness of the assessment for all groups of candidates and documents the analysis process, findings, and action taken.

1(h) In designing assessment administration procedures, the sponsor includes administrative accommodations that preserve assessment validity while addressing issues of access for candidates with disabilities.

1(i) In the course of developing or adopting a passing standard that is demonstrably equivalent to or more rigorous than the State recommended standard, the sponsor secures and reflects on the considered judgments of teachers, the supervisors of teachers, the support providers of new teachers, and other preparers of teachers regarding necessary and acceptable levels of proficiency on the part of entry-level teachers. The sponsor periodically reconsiders the reasonableness of the scoring scales and established passing standard.

**Assessment Design Standard 2: Assessment Designed for Reliability and Fairness**
(Assessment Design Standard 2 Applies to Programs that Request Approval of Alternative Assessments)

The sponsor of the professional teacher preparation program requests approval of an assessment that will yield, in relation to the key aspects of the major domains of the TPEs, enough collective evidence of each candidate's pedagogical performance to serve as an adequate basis to judge the candidate's general pedagogical competence for a Preliminary Teaching Credential. The sponsor carefully monitors assessment development to ensure consistency with the stated purpose of the assessment. The Teaching Performance Assessment includes a comprehensive program to train and re-train assessors. The sponsor periodically evaluates assessment design to ensure equitable treatment of candidates. The assessment design and its implementation contribute to local and statewide consistency in the assessment of teaching competence.

**Required Elements for Assessment Design Standard 2: Assessment Designed for Reliability and Fairness**

2(a)   In relation to the key aspects of the major domains of the TPEs, the pedagogical assessment tasks and the associated directions to candidates are designed to yield enough evidence for an overall judgment of each candidate's pedagogical qualifications for a Preliminary Teaching Credential. The program sponsor will document sufficiency of candidate performance evidence through thorough field-testing of pedagogical tasks, scoring scales, and directions to candidates.

2(b)   Pedagogical assessment tasks and scoring scales are extensively field-tested in practice before being used operationally in the Teaching Performance Assessment. The sponsor of the program evaluates the field-test results thoroughly and documents the field-test design, participation, methods, results and interpretation.

2(c)   The Teaching Performance Assessment system includes a comprehensive program to train assessors who will score candidate responses to the pedagogical assessment tasks. An assessor training pilot program demonstrates convincingly that prospective and continuing assessors gain a deep understanding of the TPEs, the pedagogical assessment tasks and the multi-level scoring scales. The training program includes task-based scoring trials in which an assessment trainer evaluates and certifies each assessor's scoring accuracy in relation to the scoring scales associated with the task. When new pedagogical tasks and scoring scales are incorporated into the assessment, the sponsor provides additional training to the assessors, as needed.

2(d)   In conjunction with the provisions of Teacher Preparation Program Standard 19, the sponsor plans and implements periodic evaluations of the assessor training program, which include systematic feedback from assessors and assessment trainers, and which lead to substantive improvements in the training as needed.

2(e)   The program sponsor requests approval of a detailed plan for the scoring of selected assessment tasks by two trained assessors for the purpose of evaluating the reliability of scorers during field-testing and operational administration of the assessment. The subsequent assignment of one or two assessors to each assessment task is based on a cautious interpretation of the ongoing evaluation findings.

2(f)   The sponsor carefully plans successive administrations of the assessment to ensure consistency in elements that contribute to the reliability of scores and the accurate determination of each candidate's passing status, including consistency in the difficulty of pedagogical assessment tasks, levels of teaching proficiency that are reflected in the multilevel scoring scales, and the overall level of performance required by the Commission's recommended passing standard on the assessment.

2(g)   The sponsor ensures equivalent scoring across successive administrations of the assessment and between the Commission's model and local assessments by: using marker performances to facilitate the training of first-time assessors and the further training of continuing assessors; monitoring and recalibrating local scoring through third party reviews of scores that have been assigned to candidate responses; and periodically studying proficiency levels reflected in the adopted passing standard.

2(h)   The sponsor investigates and documents the consistency of scores among and across assessors and across successive administrations of the assessment, with particular focus on the reliability of scores at and near the adopted passing standard. To ensure that the overall construct being assessed is cohesive, the sponsor demonstrates that scores on each pedagogical task are sufficiently correlated with overall scores on the remaining tasks in the assessment. The sponsor demonstrates that the assessment procedures, taken as a whole, maximize the accurate determination of each candidate's overall pass-fail status on the assessment.

2(i)   The sponsor's assessment design includes an appeal procedure for candidates who do not pass the assessment, including an equitable process for rescoring of evidence already submitted by an appellant candidate in the program.