Costs and Benefits of Accountability: A Case Study of Credential Candidates' Performance Assessment

Paula M. Selvester, Deborah Summers, & Edward F. Williams
California State University, Chico

Introduction

In this article we present a case study that is an exploration of credential candidates' attitudes toward a teaching performance assessment as a measure of their teaching ability, and the impact on faculty instructional decisions, practice, and attitudes toward the assessment. This study generated both qualitative and quantitative data. Results revealed that assessment, discussion, sharing results among faculty, credential candidates and partner schools can begin to open the dialogue about the presence and status of the characteristics of effective teacher preparation programs.

Theoretical Framework

In California, Senate Bill 2042 redesigned the teacher credentialing system to reflect alignment of professional standards and assessment of teachers. A teaching performance assessment was included in the new credential to judge the impact of professional preparation **(**Alpert & Mazzoni, 1998). Theorists who support standards-based, accountability systems are optimistic, offering models that predict their success (Betts

---

Paula M. Selvester, Deborah Summers, and Edward F. Williams are assistant professors in the Department of Teacher Education at California State University, Chico.

& Costrell, 2001; Haycock, 2001; Odden, 1995); however, opponents present evidence that these systems threaten to hinder authentic, student-centered learning, codify one version of knowledge, and reduce and narrow the curriculum (Amrenin & Berliner, 2002; Darling-Hammond, 1994; Ogaway, Sandholtz, Martinez-Flores, & Scribner, 2003).

The standards-based, high-stakes accountability reform movement is driven by concerns about student learning. The call is for public schools and colleges of education to improve. Odden (1995) and others (Goldberg & Traiman, 2001) frame the movement by stating that the question lies not in whether or not schools are improving, but whether or not they have improved enough to prepare students and teachers for success in a rapidly changing world. In an analysis and explanation of the function of standards-based, high-stakes accountability reform, Betts and Costrell (2001) and Odden, (1995) use microeconomics to distinguish the possible costs and benefits of standards-based, high-stakes testing policies. In their analyses, they describe a system by which students, parents, teachers, and administrators performing under a standards-based system will alter their behavior in order to be classified as winners and avoid being classified as losers, behavior they insist will advance learning.

However, other researchers maintain that there is no convincing evidence that high-stakes, standardized testing policies improve student performance (Amrenin & Berliner, 2002) or that paper and pencil tests can be used reliably to evaluate teacher performance (Popham, 1990) or student learning (Popham, 2001). Research on effective teacher preparation has emphasized the positive effect of well-defined standards supported by a shared and common vision of good teaching that are enacted consistently in both coursework and field work (Darling-Hammond, 1999).

In a study of teacher preparation programs, Delandshere and Arens (2001) examined two states, Indiana and Vermont, which had taken different approaches to teaching standards and licensure reform. Interviews with education faculty members about the meaning of teaching, the teachers' role, and their vision of schooling and education revealed that teacher educators' degree of participation in standards-based reform was influenced by their conception of teaching, learning, and schooling, and its purpose. Findings suggested that teacher educators whose vision and ideas about education and its purpose were vague or unspecified were more likely to uncritically accept the authority of and participate in standards-based reform.

Strong programs of teacher education require teacher educators' participation in the important work of developing a shared vision of effective instruction, well-defined standards, and a coherent system of assessment of course and field work (Darling-Hammond, et al., 2005).

The claims of the Delandshere & Arens (2001) study support the notion that the authority of standards and the standardization of teacher preparation without a shared, common vision of what characterizes good teacher education may lead teacher educators to adopt a narrow or reductionist approach to the complex challenge of teaching and learning.

Method

In response to the recommendations by the California Commission on Teacher Credentialing, the Department of Education at a California State University designed and implemented a Teaching Performance Assessment (TPA) as a requirement for receiving a preliminary credential. This case study was designed to investigate the impact of a standards-based teacher performance assessment on the teacher education faculty and the credential candidates who took it. Both quantitative and qualitative data were generated. Using grounded theory (Glaser, 2005), qualitative analysis led to themes that emerged from multiple readings of the notes taken from the open-ended questionnaires and focus group notes. Quantitative analysis allowed for means to be identified from data generated by surveys composed of Likert-style items and the teaching performance assessment scores.

Sample

This study was conducted on the campus of a medium-sized rural state university located in Northern California. The participants were fifth-year pre-service credential candidates (n=151) and education faculty (n=16) in a graduate teacher preparation program. The Department of Education at this university prepares up to 400 to 500 credential candidates per year and places them in schools for teaching experience in a wide surrounding area that reaches from Chico to the Nevada border to the east, the Oregon border to the north, the Pacific coastline area to the west, and as far south into the Central Valley as Fresno.

Timeline

During the redesign phase of the California teaching credential, Senate Bill 2042, the Department of Education at a California State University in Northern California began development of a teaching performance assessment (TPA), based on the Educational Testing Service (ETS) assessment model prototype, developed for the California Commission on Teacher Credentialing.

To establish validity before administering the first version of this

experimental performance assessment, the department held a conference day, inviting all cooperating teachers, supervisors and faculty to review the assessment and evaluate it. Cooperating teachers, faculty, and university supervisors (n=178) were surveyed to determine the extent to which the TPA measured the TPEs. The results of the data informed revisions of the draft and the pilot TPA was given to all of the credential candidates in both the multiple and single subject credential program (n=165) in the last semester of their fifth year teacher preparation program in the spring of 2003.

In the fall of 2003, the TPA was administered again, a follow-up candidate questionnaire was administered, and a focus group was conducted. In spring 2004, the TPA was administered, and in fall 2004, the faculty completed a questionnaire. Preliminary results were presented at the California Council of Teacher Education in October 2004 and final results were presented at the American Educational Research Association in 2005.

Data Source

Teaching performance assessment. The TPA is a set of performance tasks that are designed to measure the California Teaching Performance Expectations (TPEs). The TPEs describe what California teachers should know and be able to do upon completion of their teacher education program and prior to receiving their preliminary credential. This performance assessment was made up of two specific tasks that require credential candidates to write about how they planned, taught, assessed, and reflected upon two lessons at two intervals during a fifteen week practicum. Each task was broken down into steps that were designed to measure a range of the 13 teacher performance expectations that fall into the six broad domains of the California Standards of the Teaching Profession (CCTC, 1997). While both tasks require candidates to implement a lesson, Task B differs from Task A in that more emphasis is placed on the candidate's ability to interpret and use assessment data of student work in order to further and inform future instruction.

To demonstrate their abilities to differentiate instruction using theory and practice learned in their course work, the TPA required credential candidates to select three focus students (an English language learner, a student identified with a special need, and a student who represents a different instructional challenge) to specifically follow through two specific lesson plans taught during their second placement. In the planning stage of a lesson, for example, credential candidates must describe each focus student's characteristics regarding gender and age,

Table 1
Differences between TPA Task A and Task B

TASK A		TASK B	
Step	**Description**	**Step**	**Description**
Step 1 Reflect on practicum.	Write (or revise) a reflection on Teaching Practicum I.	**Step 1** Information About Your Class	Provide a copy of your response to Task A, Steps 2 and 3.
Step 2 Learning about Students	Select a class; collect and report information on instructional context and the class as a whole.	**Step 2** Update on Learning Environment	Report on the steps you have taken toward a positive democratic learning environment and their effectiveness so far.
Step 3 Democratic Learning Environment	Describe how to establish and sustain a democratic learning environment for your class. Write adaptations for each focus student to the target lesson.	**Step 3** Academic Instructional Planning	Select a content area, the related state-adopted K-12 content standards for students, a unit of study, and at least a three-lesson learning segment. Report on your plan for instruction for the whole class.
Step 4 Lesson Adaptations for Focus Students	Select three focus students; describe each focus student.	**Step 4** Assessment Planning	Report on your plan for assessing student learning in your target lesson.
Step 5 Instructional Planning	Select a content area, related state-adopted content standards, a unit and a three-lesson learning segment. Describe your instructional planning.	**Step 5** Lesson Adaptations for the Three Focus Students	Write lesson adaptations for each focus student.
Step 6 Analyze the Lesson	Score evidence of student learning; analyze student work to identify patterns in student understanding. Submit copies of student work samples.	**Step 6** Assessment Adaptations for Three Focus Students	Describe assessment adaptations for each focus student.
Step 7 Reflection on Instruction	If you were to teach this lesson again to the same students, what would you do similarly and differently?	**Step 7** Analyze the Lesson Through Student Assessment	Score evidence of student learning; analyze to identify patterns in student understanding. Submit samples of student work.
		Step 8 Analyze the Lesson Through Focus Student Assessment	Analyze evidence of focus student learning to identify progress toward goals and effectiveness of adaptations. Submit focus students' work samples.
		Step 9 Reflection on Assessment	If you were to teach this learning segment and lesson again, and conduct the assessment, what would you do similarly and differently?
		Step 10 Reflection on Instruction and Professionalism	Consider your experience in Practicum II; write narrative reflections to questions.

and linguistic background. In addition, they are required to discuss the students' academic language abilities, and cognitive, physical, and social and emotional development related to the subject matter of the lesson.

The candidate must address the students' cultural background, including family and home, and take into consideration health issues, and interests and aspirations related to the content area. Credential candidates also include information describing the learning environment, the objectives of the lesson, the instructional strategies, and other aspects of instructional planning. Finally, candidates teach the lesson, analyze, and reflect upon it.

The rubric used to evaluate and score the TPA responses utilizes a 5-point scale: 5 to 4 is target, 3 to 2 is acceptable, and 1 to 0 is unacceptable. Each step is given a score of up to 5 points. Each candidate receives a total score and an averaged score. A passing score must be an average of 2.5 with no scores of 0 or 1 on any step. A perfect score is 35 points. Scoring and submission are anonymous.

During this study, in order to assure inter-rater reliability, faculty and university supervisors attended a day-long scoring training. Sample TPA responses were analyzed and scored. Scores among faculty were compared and discussed to test the efficacy of the scoring rubric. The first of two tasks, Task A, was administered to credential candidates who completed it and submitted it to the department in five weeks. Teams of trained scorers were each given up to fifteen tasks to score anonymously. Inter-rater reliability was assessed using Pearson product-moment correlation coefficient, $r = .87$. Mean scores for each of the seven steps of Task A were calculated and overall mean scores were also generated. Credential candidates who failed the test were permitted to resubmit their responses. Resubmissions were scored again separately. Students who failed the second attempt were not recommended to continue in their teaching practicum and were required to retake coursework.

Credential candidate follow-up questionnaire. To collect information about the impact of the TPA on the credential candidates' learning experience, a follow-up questionnaire was sent to candidates (n=151) after they had received their final scores on the TPA. A total of 50 questionnaires were returned. The questionnaire included seven questions (Likert-style, 5 point scale) as well as two open-ended questions. This questionnaire asked credential candidates to rate the effectiveness of the TPA in contributing to their growth of knowledge, skills, and abilities as teachers. They were asked if it measured their ability to learn about the educational needs of individual students, to plan to assess students and determine growth, and to identify specific data sources for gathering information about student needs. They were also asked how well they thought the TPA was an accurate "snapshot" of their teaching practice. The overall means for Likert-style items were determined and in an analysis of the qualitative data generated from the open-ended questions, themes were coded and identified.

Focus group. A focus group was also conducted with credential candidates (*n*=10) to find out how they viewed their experiences taking the TPA. They were asked about their perceptions of the TPA, and its efficacy as a source of information about their teaching performance and as a tool for reflection and growth as a new teacher. Notes from the focus group were analyzed qualitatively for emerging themes.

Faculty questionnaire. Teacher education program faculty, full-time, tenured and tenure-track, (n=16) were surveyed with an open-ended questionnaire. At the time they were surveyed, each had read an average of 28 assessments. Survey questions asked faculty to describe specific ways it had or had not impacted their beliefs, instructional or supervision practice, and/or classroom content. Those who believed the TPA had not been useful in informing their teaching practice were asked to describe or list reasons why it was not useful. Faculty were also asked if they believed the TPA gave them information about how well the program was doing overall to prepare new teachers. Finally, they were asked how well they thought the TPA assessed what new teachers need to know and do to be successful. Qualitative analysis led to themes that emerged from multiple readings of the notes taken from the open-ended questionnaires and focus group. Quantitative analysis allowed for means to be identified for each of the seven Likert-type questions using the five point scale (5 being highest).

Results

The purpose of this study was to investigate the impact of a standards-based teacher performance assessment on the teacher education faculty and the credential candidates who took it. There is a body of research developed around economic models that predict the success of standards and testing to improve student and teacher learning (Betts & Costrell, 2001; Odden, 1995). Yet, there is research that indicates that in the absence of clearly articulated ideas about the purpose of education, faculty will accept and participate in standards-based reform (Delandshere & Arens, 2001) and narrow the curriculum and their instructional strategies to accommodate the standards (Ogaway, et al., 2003). How did the TPA impact the credential candidates? How did this department's use of a teacher performance assessment inform its teacher preparation faculty instructional practice? Had reading the completedTPA tasks impacted the faculty's curriculum? What potential did the TPA hold for program improvement?

TPA Results

The results of the TPA showed that the faculty rated the candidates'

abilities highest in the areas of learning about their students (3.31), analyzing their lesson plans (3.18), and reflecting upon their instruction (3.30). Lower scores were given in the areas of reflecting upon previous student teaching experience (2.90), establishing a democratic environment (2.82), and the lowest scores given for teaching the lesson (2.80). The overall mean score was 3.00.

TPA Impact

Quantitative data revealed that the majority of credential candidates rated the TPA a "3" (on a five point scale, 5 = significantly affecting growth) on its positive effect on their growth, knowledge, abilities, and skills as teachers. The lowest overall means were 2.6 (a measure of planning for instruction) and 2.5 (a snapshot of your teaching) respectively. In an analysis of the post-test survey and the focus group follow-up, two themes emerged: candidates consistently asserted that the work of reflection in the assessment was redundant and that they needed more support in doing the TPA from their supervisors, their cooperating teachers, and to a lesser degree, from their professors.

Redundacy. The TPA required candidates to report on the planning, teaching, and assessment of two lessons, reflecting extensively and in

Table 2
Means and Standard Deviations
of Teacher Performance Assessment (TPA) Tasks

Step	**n**	**Mean**	**SD**
1. Reflect upon previous student teaching experience	151	2.90	1.32
2. Learning about students	151	3.31	1.02
3. Establishing a democratic environment	151	2.82	1.01
4. Lesson adaptations for three focus students	151	2.97	1.17
5. Planning for instruction	151	2.80	1.16
6. Analyze the lesson	151	3.18	1.17
7. Reflect upon instruction	151	3.30	1.13
Overall Mean Score	151	3.00	

Table 3
Means and Standard Deviations of Credential Candidate Survey

Survey Question	**n**	**Mean**	**SD**
1. Learn about students	50	3.00	1.29
2. Plan for individuals	50	3.12	1.30
3. Plan for instruction	50	2.68	1.38
4. Assess for growth	50	3.00	1.11
5. Analyze and reflect	50	3.23	1.28
6. Difficulty of TPA	50	3.36	1.14
7. Represents teaching expertise	50	2.57	1.11
Overall mean score	50	3.00	

detail during each stage of the lesson. However, one of the lesson tasks focused intensely on planning and analysis of students' needs and the other lesson task focused in detail on assessment. Both tasks asked candidates to reflect after their analysis of student results, although these tasks were written for different purposes. When asked for suggestions for improvement of the TPA, one candidate's complaint, "the TPA was extremely repetitive! Reflect, analyze, reflect, analyze, reflect, analyze for 20 plus pages, come on," was indicative of respondents' perception of the redundancy both within and between tasks. The majority of the written responses (fifty-two percent) indicated that the credential candidates were frustrated with redundancy. In another characteristic response, a credential candidate advised, "The questioning in the TPA is far too redundant. The TPA should be reduced; portfolios would be more useful and show the range of what we do as teachers."

Need for more support. Requests for more assistance characterized 35 percent of candidate responses. The kinds of support candidates suggested included: (1) additional coursework, (2) closer fieldwork supervision during the task, (3) clear directions, and (4) more exemplars. Many suggested they would have been more successful or found the task more

enriching had they been able to work more closely on the TPA with their supervisors and their cooperating teachers. They also indicated that having the TPA tasks built into their coursework would have been more valuable to them as a learning experience. A typical response suggested, "I would have liked more supervisory scaffolding or help making explicit what TPA is asking for… (I was) thrilled to have successfully negotiated it!"

Faculty Impact

Faculty responses indicated that reading the TPA tasks impacted their curriculum, their teaching, and their supervision. Targeting standards, other curricular information, or specific skills faculty perceived important for success on the TPA was an important theme that emerged from faculty responses. For example, faculty reported they began "hitting harder on differentiated instruction, more specifically the English Language Development standards." A common response indicated a heightened faculty awareness of TPA content. Typical responses indicated that faculty began to make "verbally explicit links in class by stating that something would be valuable on the TPA" and "or redesigned assignments to be TPA like." Some faculty indicated that the TPA made them "more aware of their assignments" in relationship to tasks required on the TPA.

Program Impact

Did the TPA provide a source of feedback to these teacher educators about the strength of their teacher education program in preparing credential candidates? The results of the open-ended questionnaire for faculty indicated they agreed overall that the TPA revealed systematic strengths and areas for improvement.

Strengths revealed in the TPA responses were described as beliefs about the success of the teacher education program. Typical faculty responses indicated that because "it (the assessment) is closely aligned with teacher performance expectations and *our* program, it gives a good deal of information about candidates reaction and relationship to the content of our program." Some faculty thought it was a good general view of candidate performance, writing that, "it could paint an overall picture of candidate performance." Still others confirmed that it could even be a "stand alone" measure of teaching or program success.

Focus areas for improvement in the TPA were described in the area of diversity and equity. For example, faculty expressed fears that the assessment did not adequately give information about "students ability to teach in an equitable manner," or "how their pedagogy is addressing the issues of inequity." Faculty indicated that this assessment "addresses the

nuts and bolts but not whether or not we have a truly democratic, progressive, critical program." Other typical responses were characterized by such concerns as, "I think we strive to encourage our students to be reflective practitioners, but it takes more than that to be successful." In particular, faculty were dubious about the relative gains offered by the TPA compared to the time they spent on reading and scoring it, the time spent by credential candidates and whether or not it should function as a high-stakes assessment. Finally, faculty agreed overall that this performance assessment needed triangulation with other sources to be a powerful assessment of what teachers need to know and be able to do to be successful.

Discussion

There were costs and benefits for credential candidates and faculty in implementing a teaching performance assessment. The quantitative data revealed the majority of credential candidates rated the TPA as a benefit to affecting their growth, knowledge, abilities, and skills as teachers. In general, candidates perceived the process of getting to know individual student, learning and planning for their specific needs, assessing their growth and analyzing instruction, as a valuable step in their development as teachers. However, their narrative comments about their experience revealed a cost, a pervasive thread of discontent, with the requirement in the TPA to reflect at each stage of the process.

Cochran-Smith and Lytle (1999) have emphasized the importance of identifying particular kinds of knowledge development: knowledge *for* practice, knowledge *in* practice, and knowledge *of* practice. The TPA used in this study sought to elicit knowledge from credential candidates as expressed in their reflective narratives of their classroom teaching. The intent of the TPA was to elicit systematic inquiries from the credential candidates about their students, instructional planning, curriculum, and professional development.

Credential candidates were asked to provide reflections that would illustrate their knowledge, growth, and learning based upon their knowledge of subject matter, pedagogy, theories of learning and development, and teaching strategies. The faculty had intended each prompt to serve a unique purpose; credential candidates were asked to reflect upon previous fieldwork performance as it informed their current practice, reflect upon the impact of their instruction, and reflect upon the entire experience to inform their professional growth.

However, the results showed candidates viewed multiple prompts to reflect as redundant. In contrast to faculty expectations, candidates may

not have viewed reflection as a conceptual tool (Grossman, et al, 1999), a valuable means by which they could improve their understanding of their teaching. Their responses suggested that reflection was unproductive or artificial as a mode of inquiry at each stage of the teaching cycle. This attitude or belief may also reflect these credential candidates' disposition toward an "inquiry stance" (Cochran-Smith & Lytle, 1999, p.250). Although the credential candidates perceived reflection to be a cost to learning, faculty believed it to be a benefit, rating credential candidates highest in these areas, thereby revealing their beliefs in the benefit of reflection during the teaching and learning process.

Both of these explanations point toward the potential for teacher performance assessment to benefit teacher education program improvement. We want to ensure that during the course of their program, candidates participate in a coherent, purposeful cycle of learning that returns to central ideas and concepts, such as reflection, that are applied in multiple ways and within different content areas and contexts. Howey and Zimpler (1989) found in their study of strong teacher preparation programs that "key dispositional attitudes and behaviors are enabled and monitored in structured experiences" (p. 242).

Credential candidates' responses about the redundancy of reflection pointed to a gap that might not have been as apparent at the course level. Because short term interventions such as those that happen at the course level are not as effective as maintaining a consistent and unified focus across coursework (Gore & Zeichner, 1991), these responses heightened our interest in investigating more closely faculty beliefs about the role of the teacher, our conceptions about teaching and learning, and our department vision for teacher preparation. As a benefit, the study results pointed out the need for better articulation and collaboration among faculty across courses to improve program quality.

In addition to articulating and collaborating within our department, the results showed we also needed to extend articulation with our public school partners. Credential candidates' requests for more support suggested that we needed to continue to work closely with each other at the university and our partners in the schools, and to align our beliefs about instructional practices and teaching and learning into a more powerful learning community. The work of identifying children's needs, planning, teaching, and reflecting upon the results of the lesson is most productive in dialogue and discussion that is ongoing, within the university classroom and the student teaching classroom. We recognized that our credential candidates needed more opportunities to share norms and practices, such as those that were assessed on the TPA, within a structured and purposeful learning community. Opportunities to learn

about students, their backgrounds, their families, language, culture and communities are critical to teacher development.

Teaching for culturally and linguistically diverse classrooms is vital in public school classrooms where the majority of the nation's teaching population is European Americans from middle class backgrounds, yet almost half of public school children are students of color (National Center for Education Statistics, 2003). Credential candidates who develop cultural frames of reference to understand their students' needs and points of view are able to plan and teach a diverse population more successfully. Faculty responses revealed their concern that demonstrating evidence of a culturally responsive teaching practice (Gay, 2000; 2002) may not be possible in the context of a performance assessment that relies on written documentation.

As required by the TPA, the credential candidates wrote reflections upon two lessons (one after the first weeks of student teaching and the second in the completion of the tenth week) with culturally and linguistically different children, submitted student work samples, and lesson plans. Without the submission of "real time" observations, however, from cooperating teachers or supervisors or possibly even peers, faculty seemed to believe that these documents could not capture the way in which the credential candidate's pedagogy changes, and shifts in response to a growing understanding of self, students, and community. Triangulating information gathered from those participating within the credential candidates' teaching and professional learning communities is necessary to strengthen the need for multiple perspectives on credential candidates' growth. Recognizing this need to strengthen articulation of curriculum with university supervisors, cooperating teachers, and faculty was an additional benefit of this study.

Finally, there are practical costs in faculty time when implementing a teacher performance assessment. Reading and scoring teacher performance assessments during budget cuts at state institutions where faculty responsibilities have grown but support resources have dwindled are real resource concerns. Almost one half of the teacher-education faculty in this department were working towards tenure. The pressure to publish while carrying an annual teaching load of eighteen units and six units of committee work, plus the driving time for school supervision may have contributed to the faculty attitudes toward the time investment. Although scoring was "voluntary" at the time, junior faculty often feel obligated to participate to show their commitment to their departments' goals. In the most recent scoring process, in the spring of 2005, scorers were offered a modest monetary sum to compensate for their time.

Conclusion

In an era of accountability and increasing calls for alternatives to traditional certification, it is critical that state universities demonstrate their ability to implement standards-based assessment that is rigorous yet flexible, and contributes to overall program quality. There is evidence to support that strong teacher preparation programs have the following features: a shared vision of good teaching, well-defined standards of practice and performance, extended clinical experiences, strong relationships between the university and partner-schools, and extensive use of case study methods, teacher research, performance assessments and portfolio examinations that tie teacher learning at the university with their teaching practice in the schools (Darling-Hammond, et al, 2005). Assessment, discussion, sharing results among faculty, credential candidates, and partner schools can begin to open the dialogue about the presence and status of the characteristics of effective teacher preparation programs. Assessment and data collection over time can benefit all stakeholders in the education of credential candidates, providing a beneficial source information to continue the process of program self-study and improvement.

Teacher education programs can benefit by using the results of teaching performance assessments to evaluate their progress toward preparing effective teachers. Triangulation of information from supervisors and cooperating teachers will improve the reliability of the TPA as a source of feedback about candidate pedagogy enacted in the classroom. Above all, attaching high stakes to these assessments is a cost that threatens to compromise the purpose of gathering data from which to improve programs and curricula, and serve to suppress faculty and student dialogue about its purpose and value.

The TPA can also provide a rich context to open dialogue among faculty that goes beyond the daily course level discussions to a focus on the purposes, values, vision, and mission of their teacher education program. Teacher education faculty committed to questioning their own beliefs about standards and assessment can benefit from using the standards as guidelines while avoiding the cost of narrowing their curriculum to target state standards that may not include all of the knowledge, skills, and abilities their professional opinions support.

References

Ahlquist, R. (2003). Challenges to academic freedom: California teacher educators mobilize to resist state-mandated control of the curriculum. *Teacher Education Quarterly,* 30(1), 57-64.

Alpert, Senator & Massoni, Senator. (1998). Senate Bill 2042, California State Legislature. Retrieved from http://www.Leginfo.ca.gov/pub/97-98/bill/sen/sb-2001-2050 on August 1, 1003.

Amrenin, A. L., & Berliner, D. (2002, December). *The impact of high-stakes tests on student academic performance: An analysis of NAEP results in states with high-stakes tests and ACT, SAT, and AP test results in states with high school graduation exams.* Tempe, AZ: Education Policy Studies Laboratory, EPSL-0211-126-EPRU, http:// edpolicylab.org.

Betts, J. R., & Costrell, R. M. (2001). Incentives and equity under standards-based reform. In D. Ravitch (Ed.), *Papers on education policy 2001* (pp 9-56).Washington, DC: Brookings Institution Press.

Business Round Table (1996). *A business leaders' guide to academic standards.* (*Web* Pagebrtable.org/pdf/80.pdf): Business Round Table.

California Commission on Teacher Credentialing. (1997). *The California standards for the teaching profession.* Sacramento, CA: Author.

Cochran-Smith, M., & Lytle, S. (1999). The teacher research movement: A decade later. *Educational Researcher*, 28 (7), 15-25.

Darling-Hammond, L. (1994, August). National standards and assessments: Will they improve education? *American Journal of Education, 102*, 478-509.

Darling-Hammond, L. & Ancess. J. & Falk, B. (1995) *Authentic assessment in action: Studies of schools and students at work*. New York: Teachers College Press.

Darling-Hammond, L. (1999). Educating teachers for the next century: Rethinking practice and policy. In G. A. Griffin (ed.), *The education of teachers* (pp.221-256). Chicago: University of Chicago Press.

Darling-Hammond, L. (2000). *Teacher quality and student achievement: A review of state policy evidence.* Educational Policy Analysis Archive, 8(1), 35 pages retrieved online from http: //epaa.asu/epaa/v8n1/.

Darling-Hammond, L., Hammerness, K., Grossman, P., Rust, F., Shulman, L. (2005). The design of teacher education program. In L. Darling-Hammond and J. Bransford (Eds.), *Preparing teachers for a changing world* (pp. 390-441). New York: Teachers College Press.

Delandshere, G., & Arens, S. A. (2001). Representations of teaching and standards-based reform: Are we closing the debate about teacher education? *Teaching and Teacher Education, 17*, 547-566.

Gore, J. M., and Zeichner, K.M. (1991). Action research and reflective teaching in preservice teacher education: A case study from the United States. *Teaching and Teacher Education, 7*(2). *119-136.*

Gay, G. (2000). *Culturally responsive teaching: Theory, research, and practice.* New York: Teachers College Press.

Gay, G. (2002). Preparing for culturally responsive teaching. *Journal of Teacher Education,* 53 (2), 106-116.

Goldberg, M., & Traiman, S. L. (2001). Why business backs standards. In D. Ravitch (Ed.), *Brookings Papers on Education Policy, 2001*. pp. 75-130. Washington, DC: The Brookings Institution.

Glaser, P. (2005). *Grounded theory perspective III*. Mill Valley, CA: Sociology Press.

Grossman, P. L., Tompson, C., & Valencia, S.W. (2002). Focusing the concerns of

new teachers. The district as teacher educator. (see 515 in Preparing Teachers.)

Haycock, K. (2001). Closing the achievement gap. *Education Leadership, 58*(6), 1-7.

Holmes Group. (1996). *Tomorrow's teachers.* East Lansing: MI. Holmes Group.

Howey, K. R. & Zimpher, N. L. (1989). *Profiles of preservice teacher education: inquiry into the nature of programs.* Albany, NY: State University of New York Press.

National Center for Education Statistics. (2003). *The condition of education 2003.* Washington, DC: U.S. Department of Education.

Nelson, J. (2003). Academic freedom, institutional integrity and teacher education. *Teacher Education Quarterly, 30*(1), 65-72.

Odden, A. (1995). *Educational leadership for America's schools.* New York: McGraw-Hill.

Ogaway, R., Sandholtz, J. H., Martinez-Flores, M., & Scribner, S. P. (2003, Spring). The substantive and symbolic consequences of a district's standards-based curriculum. *American Educational Research Journal, 40*(1), 147-176.

Perlman, C. (2002). An introduction to performance assessment scoring rubrics. In C. Boston's (Eds.), *Understanding scoring rubrics* (pp. 5-13). College Park, MD: ERIC Clearinghouse on Assessment and Evaluation.

Popham, W.J. (1990). Face validity: Siren song for teacher tests. In J.V. Mitchell, S. Wise, & B. Plake (Eds), *Assessment of teaching: Purposes, practices, and implications for the profession* (pp.1-4) Hillsdale, NJ: Lawrence Erlbaum Associates.

Popham, J.W. (2001). *The truth about testing.* Alexandria, VA: Association for Supervision and Curriculum Development.

Ravitch, D. (2001). The case for national standards and assessments. *The Clearing House, 69*, 134-136.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.

Sleeter, C. (2003). Reform and control: An analysis of SB 2042. *Teacher Education Quarterly, 30*(1), 19-30.

Tellez, K. (2003). Three themes on standards in teacher education: Legislative expediency, the role of external review and test bias in the assessment of pedagogical knowledge. *Teacher Education Quarterly, 30*(1), 9-18.

U.S. Department of Education. (2002). *Meeting the highly qualified teachers challenge: The secretary's annual report on teacher quality.* Washington, DC: Office of Postsecondary Education.

Wolters, C., Pintrich, P., & Karabenick, S. (2003, March). Assessing academic self-regulated learning. Paper prepared for the Conference on Indicators of Positive Development: Definitions, Measures, and Prospective Validity. Sponsored by Child Trends, National Institutes of Health. 1-49.

Zimmerman, B. J. (2001). Theories of self-regulated learning and academic achievement: An overview and analysis. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (pp. 1-38). Mahwah, NJ: Lawrence Erlbaum Associates.