

Restoring Balance: A Chronology of the Development and Uses of the California Standards for the Teaching Profession

**By Andrea Whittaker,
Jon Snyder, & Susan Freeman**

Andrea Whittaker is a professor in the College of Education at San Jose State University, San Jose, California; Jon Snyder is a professor in the Graduate School of Education at the University of California, Santa Barbara; and Susan Freeman is co-director of the Central California Writing Project at the University of California, Santa Cruz. The authors contributed equally to the development of this article, and names are listed in reverse alphabetical order.

Introduction

In January 1997, after four years of development and study, the California Commission on Teacher Credentialing voted unanimously to adopt the *California Standards for the Teaching Profession* (CSTP). In the informational report to the Commission the staff wrote,

In the 150-year history of California education, the *California Standards for the Teaching Profession* are the first teaching performance standards that have statewide validity.... The governance of teaching can assume the stature of a profession only when teaching practices are governed Professionally. Completion and adoption of the *California Standards for the Teaching Profession* are milestones in the long-term effort to foster professionalism in California teaching. (p. 18)

In the four years since their adoption, the CSTP have become the cornerstone document for teaching

Restoring Balance

policy in the state of California. They have been and are being used to (1) support participants in the Beginning Teacher Support and Assessment Program (the function which drove their initial formulation); (2) guide the re-design of preservice teacher education program standards; (3) serve as the centering constructs for the teacher performance assessment all beginning teachers in the state of California will need to pass to receive a license to teach; (4) frame the Peer Assistance Review (PAR) program, the Governor's initiative to provide a structure to support experienced teachers in improving their practice or remove them from the profession if they do not; and (5) frame employment evaluation processes in many districts throughout the state.

Given the centrality of the CSTP in teaching policy in California, this article offers an historical perspective on the origins, development, and revisions of the Standards. In particular, the article focuses on the issues, concerns, and dilemmas that shaped the Standards in each of their developmental phases. Often ignored in the press of public policy is the constant need to evaluate the original intents of standards and to assess their current uses in light of that history. This taking stock is a necessary stage in the evolution of standards, particularly when assessments are designed around those standards and high stakes attached to the assessments. At a time when accountability is demanded of educators across the nation, it is vital that policy makers hold themselves accountable as well for policies and practices affecting thousands of professional educators and the access to quality teaching by millions of students.

The three authors bring a unique perspective to the evolutionary development of the CSTP. Between us, we, with many other educators, psychometricians, politicians, and government officials, have been involved at every phase of their development at all levels of the system. As professional educators, we helped craft them conceptually as well as helped to write them at each stage. As teacher educators, we have used them to support and assess preservice and inservice teachers. As researchers, we were members of the team that conducted a study of their validity and revised them based upon the results of that study. While we deserve neither credit or blame for the standards, we certainly possess an "insider's perspective" on their development. In order to check our own memories and biases, we based our analysis on a review of existing public documentation of the Standards development process as well as minutes and notes from meetings at which we were present.

Given our roles in the development and adoption of the Standards, we clearly are not in the business of bashing either the Standards or their developers. In one way, our goal for ourselves in writing this article is the same as we have for the reader: to assess and support our own uses of the Standards. In another, our goal is to offer a cautionary tale for educators who, more and more often in today's world, find themselves caught in the rub of public policy and professional practice.

What follows is a chronology of the development of the CSTP and the contexts in which they were crafted. We seek to let this history bring to the surface the issues

that influenced, and still influence, the uses and consequences of the California Standards for the Teaching Profession.

Early Development of the Standards—1991-1994

From 1988 to 1992, the work of the California New Teacher Project (CNTP) revealed the need for a commonly understood set of expectations about the knowledge, skills and abilities required by beginning teachers. In a summary evaluation of its efforts in beginning teacher support, the CNTP recommended that:

A state framework must be developed that will outline the knowledge, skills, and abilities expected of beginning teachers, and will serve as the basis for accurate information about their performances. ... The framework must recognize that teaching is complex, that teachers need time to develop teaching expertise, and that a teacher's decisions, practices and perspectives are interrelated. The framework must encompass the most significant elements of teaching, but must not be a checklist of teacher behaviors. (Charge to the Task Force: Development of the Framework of Knowledge, Skills, and Abilities for Beginning Teachers in California." Commission on Teacher Credentialing and California Department of Education. March 1994. p.1.)

In 1991 (see timeline in Figure 1), the California Department of Education (CDE) and the California Commission on Teacher Credentialing (CCTC) awarded a contract to the Far West Laboratory for Educational Research and Development (FWL, now WestEd) to prepare a draft framework. Based on an extensive review of the literature and conversations with national leaders in beginning teacher support and assessment, the FWL draft framework organized teachers' essential knowledge, skills, and abilities into six domains based primarily on the teaching effectiveness literature. The document was considered a "work in progress" at the time the CNTP funding ended in 1992.

The final report of the CNTP emphasized the need to inform support of new teachers with authentic assessments of their teaching practice. The CNTP call for a state framework for beginning teachers was further developed through the SB 1422 legislation passed in 1992. SB 1422 also initiated the Beginning Teacher Support and Assessment (BTSA) Program to phase-in support and assessment services to all new teachers in California. SB 1422 provided the legal basis for the CCTC and the state superintendent of schools to "develop and adopt a broad framework of challenging, realistic expectations for beginning teachers," and to disseminate the Draft Framework widely to improve teacher education and support programs.

As a result of SB 1422, the Beginning Teacher Support and Assessment Program (BTSA) replaced the CNTP, and a cohort of BTSA sites integrated teacher assessments into their support programs. Several BTSA programs took on the challenge of developing local formative assessment tools and were given the option of using those or the ETS Praxis (later Pathwise) assessment system which was then being introduced into BTSA. The Santa Cruz New Teacher Project (SCNTP) had designed the Con-

Restoring Balance

Figure 1
Timeline for CSTP Development and Related Statewide Initiatives

1991-1992	1994	1994-1995	1995-1996	1997	1998	1999
California New Teacher Project ends (CNTP)	Far West Lab develops Portfolio for BTSA	Statewide Taskforce led by Carol Bartell revises Domains as "Draft Framework"	Santa Cruz New Teacher Project (SCNTC) team conducts validity study of "Draft Framework"	CTC and State Board adopt the "California Standards for the Teaching Profession"	BTSA begins statewide implementation	SB2042 legislation passes, panel formed
BTSA begins with 30 pilot projects	and				CFASST widely implemented	Peer Assistance and Review (PAR) begins
Far West Lab (FWL) develops "Six Domains"	Revises Domains, adds Sub-domains and Reflective Questions	Adds narratives	Revised as the "California Standards for the Teaching Profession"	SCNTC team drafts development scales		CFASST required of most BTSA programs
SB 1422 begins	Draft developmental scales			\$110 million BTSA		
				CFASST development begins (BTSA directors, WestEd, CCTC, and ETS)		
				CCTC, WestEd and ETS develop "descriptions of practice" (DOPs)		

tinuum of Skills, Knowledge and Attitudes (Moir & Garmston, 1992) during the last days of the CNTP, and this document was widely disseminated around the state.

SB 1422 also mandated further development of the Draft Framework of Teacher Knowledge, Skills and Abilities. The "Six Domains" of the Far West Lab's Draft Framework became a central feature driving the assistance of beginning teachers. At this time, Far West Lab was awarded the contract to develop a portfolio assessment process for use by BTSA programs. As part of this portfolio development, FWL staff convened a small group of BTSA directors and other representatives to revise the "Six Domains" document for use within the formative portfolio process. The group, meeting throughout the winter and spring of 1994, tackled the issue of

specificity of the description of best teaching practices and developed the format of the six domains, elements (or sub-domains) that specified aspects or components relevant to each domain, and reflective questions that a teacher might ask about that element to inform her/his teaching. Indicator questions were phrased to read “How do I...” to show that the purpose of assessment was to inform beginning teachers’ growth. Simultaneous with the work of this group, the statewide Framework Task Force commenced its work (see detailed description below). CCTC representatives routinely passed Far West products to the task force.

Throughout the development process participants raised a number of key issues/concerns. First, the group perceived their goal as the creation of a formative model for the assessment of beginning teachers’ professional growth. Participants “did not want to develop a summative model for evaluation, but for professional growth, so that advisors can better target areas for potential growth.” One participant suggested that “the teacher-as-learner cannot be separated from the teacher-as-practitioner. I see assessment as looking at where we are and where we are going, and not as summative evaluation.” Another offered that the standards and assessment processes must be congruent with what the teacher is doing in the classroom.

CCTC participants in the group agreed that the information gathered through such a model should be formative and useful to teachers, but also insisted that it must offer a reliable measurement of teaching performance. The group agreed that such an assessment should not be used for “making a credentialing decision.” The group suggested three criteria to improve current assessment tools used in BTSA programs. They should:

- u be more structured and rigorous
- u reflect teacher performance
- u measure how a teacher is growing in the profession

Appropriate Uses of the Standards

During these discussions a tension surfaced regarding the “hand holding” nature of support programs vs. the needs of principals and others to “make judgements” about retaining teachers. Most in the group felt that the assessment tools needed to be designed to develop a beginning teacher’s ability to be self-reflective and metacognitive, and to help them identify areas for growth and self-assessment, and to “develop habits of mind and attitudes.”

A second tension related to the perceived need to redefine the domains in terms of behaviors and performance. What does it look like when teachers are actually doing what is recommended in the domain or element? This tension led to the construction of “developmental scales.” The purpose of the scales was to begin to establish exemplars that could support reasonable judgements about performance as well as to reflect the recursive and digressive nature of development. The group offered four levels of practice to define the development of teacher practice over

Restoring Balance

time (emerging, practicing, integrating, innovating). An early draft of such a scale was piloted in the initial BTSA portfolio process in summer 1994.

There was also ongoing concern that the domains “need to reflect what we want a teacher to be versus what a typical teacher is now.” The group agreed that their task was to set high expectations to drive support for all beginning teachers in the context of ongoing professional development. This is one of the reasons why the integrating level of the scales reflects a level of sophistication where an individual is enacting all domains of teaching in a fluid and integrated manner.

The group began initial conversations about how to display the domains to teachers as a non-linear schematic that demonstrates the interactive, interdependent nature of the relationship between the domains and their logical sub-domains to represent a holistic view of teaching. The group sought to make a clear picture of the complexity of teaching (and learning) for beginning teachers without overwhelming them. It was suggested that the domains and their corresponding elements be viewed as a mobile to illustrate the interactive, interdependent nature of teaching and learning. This analogy later became an important conceptual image that assisted in the validity study and final drafting of the CSTP.

Statewide Task Force—1994-95

The BTSA Program began serving new teachers in 1992. In 1994, the CCTC and the Department of Education established the Framework Task Force to undertake the revision of the Draft Framework developed by Far West Laboratory and to recommend the adoption of a completed document that would guide the work of teacher preparation programs, induction activities, and professional development for all teachers. This Task Force, led by Carol Bartell, represented a cross-section of educators selected for their knowledge and understanding of best teaching practices, teacher education, and teacher development in California. They were charged with a review of current literature; an examination of national and state professional standards that had relevance to California’s efforts; an analysis of the appropriateness of the domains and subdomains; and consultation and dialogue with other educators across the state. Although the Task Force did not take the development of the Draft Framework to state adoption, they did collaboratively produce a revised document that has been widely used in the thirty BTSA programs throughout California. Their revision, completed in August, 1995, served as the basis for the validity study of the standards described in the following section.

The Draft Framework Task Force met for nearly two years. They took the task and their role seriously. The attendance was remarkably consistent and high. With only one or two exceptions, people attended all the meetings, did all the homework between the meetings, and left their day-jobs behind them when the Task Force met so they could focus on the work at hand. Some Task Force members attributed the calibre of their efforts to a constellation of factors:

- u the exceptional facilitation provided by Carol Bartell of the CCTC staff;
- u the recognition of the significance of the work;
- u the nature of the individual professional development that was a by-product of the work (e.g., considerable reading of the literature, presentations from BTSA and teacher education practitioners, psychometricians from ETS and IHEs, and simultaneous grounding in abstract conceptual frames and concrete contexts).

As a starting point, the Draft Framework Task Force was given the domains and the very incomplete indicators originally developed by Far West Lab in 1992. There was on-going confusion, and some tension, about how much the Task Force could alter those domains. The general consensus was that in their current form, they did not capture the requisite integration of knowledge, skills, and dispositions required for quality teaching. They tended toward the atomistic rather than a holistic and developmental view of teaching. The feedback the group received on this issue varied over time. In the end, the Task Force built upon the Far West work. The six FWL domains were recognizably embedded in the revised document, and the Task Force felt comfortable that their perspective on revisions to the original draft had then been honored.

In order to assure that the group was keeping in touch with the rest of the world, the Draft Framework twice underwent extensive field review during their work together. The feedback from the field mirrored the issues and dilemmas with which the group was wrestling. These shared recurring themes included issues related to the heterogeneity of teachers and entry points into the profession; how to balance generic teaching standards with subject and context specific teaching standards; balancing educational with psychometric uses of the standards (what would be the appropriate uses of the standards?); and linking teaching standards with teacher opportunities for learning. Each of these themes is elaborated below.

Who are the Standards For?

The Draft Framework Task Force, pushed by members who worked with emergency permit teachers, often grappled with how high the standards (domains) of the Framework should go. Both ends of the spectrum were represented in the Task Force. While no-one argued for lowering standards, some felt that there were teachers in schools who would not know what these standards meant, let alone how to enact them with children. At the other end, some members of the group felt strongly that these standards should serve as an advanced vision of quality teaching—not just what could be expected of beginning teachers. The underlying forecast was that the standards drawn from the Framework would not remain strictly a BTSA document, but would function through the entire teacher development continuum from preservice, as standards for preservice programs, to teacher evaluation standards for experienced teachers in districts, as well as standards for advanced professional status. This forecast has since proved accurate.

Generic vs. Content and Context Specific

Given the time frame of the Task Force and a belief that, no matter how specific domains might become, they would still need to be flexibly understood, enacted, and evaluated in different contexts, the Task Force opted for generic standards that provided language that could be applied in all contexts of teaching. The group was aware this would require extensive professional skill on the part of support providers as well as vast changes in how preservice programs, BTSA, districts, and the state selected, trained, and assigned support providers. This is why the group argued adamantly that the standards should never be seen without accompanying program standards and support materials. The psychometricians and the subject matter specialists of the group expressed deep reservations with this perspective.

The group took a different tack regarding language and cultural context issues. The debate was whether or not to make these issues into a separate domain or to infuse them throughout the existing domains. Ultimately the task force decided in favor of the infusion model and explicitly included language dealing with specific student population issues in each domain—a decision supported by the validity study but undone following the validity study (see later chronology for details).

Appropriate Uses of the Standards

Most Task Force members foresaw the use of the Framework as standards for assessments to which high stakes would be attached. Task Force members tended towards one of three positions (sometimes simultaneously):

- u high stakes assessments are bad so these standards should not be used with them;
- u high stakes assessments are bad, but they are coming anyway, and these would be better standards to use than anything else;
- u high stakes assessments are good and we should craft these standards to maximize their use in that regard.

Discussions of such consequential validity issues occupied much of the Task Force's time. The group was repeatedly told that only professional development opportunities and support would be grounded in the standards drawn from the Draft Framework, not individual licensure decisions or employment evaluation. Presentations from psychometricians from the state and private industry, however, left a majority of the group uncertain that these standards would never be used as the foundation for high stakes assessment, an uncertainty history has proven accurate.

The group had no disagreement, however, that the fundamental purpose of these standards should be to guide and support practice. The Task Force focused its attention and decisions on educational and pedagogical concerns rather than psychometric or evaluation concerns. One example in this regard is the use of questions in the standards rather than statements. Questions, the group felt, opened practice to inquiry and supported growth, while statements cut off conversation and

tended to transform teaching into behavioristic type checklists that are quite useful for low inference judgments but much less so for understanding or supporting quality teaching.

Linking Teaching Standards with Teacher Opportunities For Learning

Given the prognostication that the state would use the standards to shape high stakes assessments, the Task Force adamantly and unanimously believed the standards should never stand alone. The standards would not be complete unless they were accompanied by a preamble, empirically-based developmental scales, program standards, and descriptions of effective, reflective practices.

The preamble the group drafted addressed several assumptions about teaching that constituted the soil from which the standards grew:

- u a view of teacher development as a complex and sophisticated ebb and flow of multiple factors rather than a chronologically defined pattern universally followed like the timing and the order of the growth of teeth;

- u a view of teaching as holistic and not atomistic thus creating the need to use the standards as a multi-dimensional whole, not discrete items in a two-dimensional linear universe.

The group believed the standards themselves would be semi-useless without empirically-based developmental scales that showed what these practices “looked like” at different levels of sophistication. The group saw program standards as a key accountability mechanism that assured teachers opportunities for learning, practicing, and assessing their development in the standards. Much like the relationship between the standards and the scales, the group felt program standards would not be particularly useful without descriptions of promising practices along the continuum of teaching. The group argued that these descriptions should serve as models for other educators to adapt and adjust to their context-specific needs and not as pre-determined molds into which to force people and programs.

Several significant changes were made to the Draft Framework between the time the Task Force completed its work and the formation of the 1995-96 validity study research team. The preamble the Task Force had developed was modified with less emphasis on: (a) the supporting documents and (b) the definitions of holistic and developmental. In addition, the new draft no longer contained much of the specific language related to context diversity (e.g., language and cultural diversity). The Draft Framework was presented without the accompanying materials. Finally, the new draft contained several new elements at the sub-domain level.

Most likely, two sources played the major role in these changes: the legislature and the BTSA Interagency Task Force (made up of CTC and CDE staff). The legislature continued to meet and address teacher standards issues during the years the Draft Framework Task Force met. Legislative actions changed the Draft

Restoring Balance

Framework directly and indirectly. Directly, they mandated that teaching standards address certain issues that received public attention at the time. Indirectly, the legislature passed SB 1422 which charged the CCTC to establish a representative panel to completely redesign the architecture of teacher preparation in the state. This group began meeting late in the work of the Draft Framework Task Force (see timeline in Figure 1). Given the panel's charge and the centrality of teaching standards in teacher education, the panel carefully analyzed the standards, making some edits, prior to recommending that the standards form the basis of all teacher education in the state. The Interagency Task Force also made changes to the standards—especially in the preamble and in the area of what supporting documents to include.

Validity Study of the Draft Framework—1995-97

The CDE and the CCTC, recognizing, in early 1995, the need to initiate a validity study of the Draft Framework, sent out a Request for Proposals for an Augmentation Grant that would, in part, fund such research. The CDE and CCTC contracted with the UC Santa Cruz New Teacher Project (SCNTP) to complete the research necessary to validate and revise the Framework. Ellen Moir, Director of the SCNTP and Teacher Education at UC, Santa Cruz, the Principle Investigator for the grant, gathered a research team, including the three authors of this article. They completed the validity study and a revision of the Draft Framework between July, 1995 and November, 1996. The new version of the Framework was titled the California Standards for the Teaching Profession (CSTP). The research team also produced a set of developmental scales to accompany the Standards in July, 1997, as part of the second phase of the validity study .

The state directed the validity study to “determine if the Draft Framework embodies and accesses the critical knowledge, skills and abilities needed by beginning teachers.” Initial meetings to organize the scope of work for the validity study and revision of the Draft Framework began in the summer of 1995. The meetings brought together members of the BTSA Interagency Task Force, research contractors from Far West Lab and Educational Testing Service (ETS), and the research team charged with validating and revising the Draft Framework. Several members of this group had been involved with the development of the Draft Framework. The state's vision of the Draft Framework and its further development under the proposed validity study was set out, and the status of the Far West Portfolio, and the ETS Praxis/ Pathwise assessment system were reviewed in the context of BTSA assessment practices.

At that time, assessment within BTSA was fragmented and inconsistent, each BTSA site following either its own structure and implementation plan or using the FWL or ETS portfolio models. The need for a coherent, consistent approach to beginning teacher assessment impelled the establishment of clear standards for the

profession, and for assessment tools and procedures that effectively and appropriately measured teachers' performance against the standards.

Members of the BTSA Interagency Task Force outlined the conceptual design for the new Standards and supporting materials, and raised concerns about the purpose and consequences of such standards. Some members of the Task Force and research team emphasized the need to retain the formative, teacher-centered nature of the BTSA assessment process. They argued that the Draft Framework should serve to promote teacher reflection, and to support and inform beginning teacher practice.

As the scope of work evolved during the summer of 1995, two separate validity studies, Phase I and II, were shaped. Phase I included two components: the development and administration of a survey and a series of focus groups to provide opportunities for more sustained and in-depth data with which to revise the standards.

Bartell, who led the Framework Task Force, maintained that "the Draft Framework itself is a vision; it doesn't define specific levels of performance. The Framework shouldn't be much more than it is now - it's a conceptual framework." In preparation for the validity study, Bartell edited and completed her group's revisions, focusing on consistent language and clarity. Bartell also edited the Task Force's introduction to the Draft Framework to provide context and clarify the purpose of the document. The validity study research team began work on the Draft Framework without receiving a copy of this introduction.

Validity Study Design

In its broadest sense, the validity study sought to determine whether the Draft Framework described what it sought to describe—a definition of teaching which was simultaneously: (a) attainable for beginning teachers in a supportive environment; (b) a vision of teaching towards which all teachers can grow; and (c) representative of the kind of teaching all students deserve and our communities require. We used a tripartite conceptualization of validity to organize the methodology of the validity study of the Draft Framework—content validity, construct validity, and consequential validity.

Closely aligned with what is traditionally called "face validity," the content validity frame focused on whether or not the document "made sense," that is, did the document, in a clear and comprehensible manner, describe the most important elements of good teaching? Was anything left out?

Construct validity focused on the deeper conceptual underpinnings of the document, whether it "hangs together" as an integrated holistic set of constructs that match the integrated holistic nature of teaching. Behind language issues and the specifics of the domains, and the sub-domains was the framework a coherent document? Was the underlying coherence accessible to the multiple roles and developmental levels of teaching of the potential users of the document?

Consequential validity focused on the potential uses of the Draft Framework. That is, how might the framework be used to support its goals? What potential uses

Restoring Balance

might inhibit its intended benefits? In short, what were the potential consequences of the Framework?

We used two data gathering techniques, a survey and focus groups. The survey was piloted by BTSA directors in December, 1995, and after minor modifications, sent to a larger sample group in late January, 1996. The focus groups were conducted during April and May, 1996.

The Survey Study

The survey consisted of 66 items including:

- u Likert scales rating the importance of each of the domains and subdomains,
- u Likert scales rating the overall coherence of the document,
- u Likert scales rating the vocabulary (clarity) of the document
- u Open-ended items requesting feedback on the completeness (what is missing?) and the potential uses of the Standards (both positive uses and fears regarding possible misuses).

It is important to reiterate that the document reviewed by survey respondents was different than that envisioned by the task force and research team. That is, the document accompanying the survey included only the “standards” as revised following the task force work and did not include the preamble or the materials the task force requested always accompany the document.

Approximately 1,100 surveys were distributed to a sample that included BTSA administrators, BTSA support providers, BTSA participants, non-BTSA beginning teachers, school site and district administrators, experienced teachers, and teacher educators. The sample was not randomly selected and was most likely skewed towards people who would be familiar with state efforts in regards to developing standards as well as working with programs that supported teacher development. We received 245 codable surveys in response—a 22 percent return rate. The response rate generated sufficient numbers of respondents from each category of respondent to ascertain any rating differences between them. The response rate was also sufficient to establish an “approval rating” for the Framework and the domains. That is, rather than seek hypothesis testing levels of significance, we sought polling levels of certainty.

The results, as with most surveys of this nature, were overwhelmingly positive. Survey respondents nearly always find it difficult to say something is NOT important. The Draft Framework received approval ratings between 93 and 99 percent.

Several concerns did emerge from the comments and open-ended questions. These concerns mirrored those raised at earlier stages of the development work and were affirmed in the focus groups later in the validity study. The most commonly voiced concern with the domains/standards was the insufficient time in the beginning teacher’s life to meet them, raising the question that perhaps the expectations were unreasonable for the induction level. The second concern had

to do with the language of the domains. There was, to these respondents, too much jargon. The third, much less prevalent, concern expressed regarded “redundancy” among the domains. Task force members attributed this concern to the missing preamble statements regarding the holistic nature and uses of teaching standards.

The survey served two purposes. The first was political in nature—to assuage fears that if the standards were approved there would be an outcry from some group that they were wrong. The 93-98 percent approval rating from the survey served that purpose. The second purpose was more conceptual and educational—to focus the more finely honed methodologies of the validity study in order to revise the domains so that they would better support beginning teachers and be more useful tools for those supporting beginning teachers.

The Focus Group Study

Four focus groups were convened in the Spring of 1996 to facilitate an in-depth examination of the Draft Framework to explore the language and structure of the document more comprehensively than the survey. The focus groups included a cross-section of educators representing preservice and induction. These represented the same constituencies as respondents to the survey. The focus groups were conducted in four different areas of the state—San Diego, Los Angeles, Sacramento, and Santa Cruz. The 73 total participants came from large and small urban areas, as well as suburban and rural areas. Participants were selected to reflect the diverse ethnic, social, and economic populations within California, including large immigrant and linguistically-diverse populations.

Focus group participants were also selected to allow for a balance between educators with a working knowledge and familiarity with the Draft Framework, and those who had no or relatively little experience with the document. Educators who were known to be interested in the induction process, and who could articulate an understanding of the essential elements of best teaching practice, were asked to join the focus groups. Beginning teachers were vital to the conversation as they could most immediately reflect the value and accuracy of the Draft Framework as a definition of teaching practice.

The focus group process was designed to examine content, construct and consequential validity issues. Focus groups engaged participants in four interrelated activities toward this end:

- u a review of the Draft Framework for clear and comprehensible language
- u an analysis of language and content in the Framework
- u visual metaphoric representations and large group discussion
- u participant reflections at the conclusion of the focus group meetings in which each participant was asked to respond to a set of two questions concerning the Framework’s most valid purposes and its inappropriate uses. They were also asked to give any final comments on the Draft Framework or the focus group process.

Restoring Balance

The results of the focus group sessions suggested revisions to the language and format of the document and the potential consequences of its use. When asked if the document described the most important elements of good teaching (content validity) focus groups confirmed that as a whole, the Framework represented essential content knowledge for teachers. This was congruent with the survey data.

Document clarity (clear and comprehensive language), however, was an issue for the focus group participants. In all four groups, participants noted the need to replace jargon with commonly understood, clear, concise words and phrases that accurately describe teaching practice. Participants, especially the more experienced support providers, maintained the need to keep some words and phrases identified as “jargon” because they felt these represented common professional language. The groups recommended that the Framework be accompanied by a glossary that defined professional terminology.

In responding to what was left out of the Framework, participants across the focus groups suggested that all levels of the Framework should provide greater attention to reflective practice, diversity, second language pedagogy, and affective learning. These suggestions were consistent with the spirit of the decision of the Draft Framework Task Force to address issues of diversity and language specifically at the element and reflective question level.

As with earlier stages of the work, the focus groups expressed concerns with how the standards should represent the complexity of teaching as a holistic and developmental endeavor. Participants urged an examination of “how the Framework defines the relationships inherent in teaching and learning,” maintaining that the document “does not show the developmental nature of teaching.” They recommended that indicators include “reflective questions on linkages between domains to help focus the interrelationships.” Participants also reported that the coherence between and within domains was also not clearly evident.

In the original Framework, the domains appeared in a numerical sequence which could be misinterpreted to mean that they should be addressed by teachers in that particular order and to imply that teaching is a set of sequenced, discrete tasks. The data from the focus groups’ large group discussions strongly suggested that the numbers be removed so that the standards reflect a more integrated, holistic view of teaching. In addition, the Draft Framework was accompanied by a one page summary matrix of the six domains and their corresponding subdomains. Designed as a grid of six numbered squares, this matrix reinforced the notion of discrete and sequential tasks. The focus group participants suggested a revised graphic representation that reflected a more holistic view.

The focus groups viewed the original title, the Draft Framework of Expectations for Beginning Teachers, as too limited. They argued the title did not define what teachers should know and be able to do throughout their careers. The text of the domains described accomplished practice and as such, defined a view of teaching that may seem overwhelming to new teachers. Therefore, the title was changed in

order to be inclusive of all teachers in California. In addition, validity study participants expressed concern over the use of the word “Framework” in the title, suggesting that it was too easily confused with the state curriculum frameworks for subject matter learning. Consistent with the terminology of national and state level initiatives, the title was changed to reflect the document’s function as standards.

Throughout the focus groups, reflection was noted as a major, if not the primary purpose of the document, and its importance in guiding beginning teacher thinking and practice was consistently mentioned. Sacramento participants stated that the “domains help the beginning teacher to find and identify evidence of positive growth” and should be used to “guide the collection of evidence and the development of new teacher portfolios.” Further evidence supporting the consequential validity of the Draft Framework was found in the participants’ belief that the document provided a common language for reflection, support and assessment for all teachers.

Focus group participants consistently expressed concerns about the misuse of the document for summative teacher evaluation. Participants warned that “using the Framework to evaluate discrete points of performance in a summative way misconstrues the purpose as a document for reflection and self-assessment.” Focus groups participants saw the document’s value in helping beginning teachers to define goals for the development of their practice, for its ability to define the scope and complexity of teaching, and to provide a set of standards for excellence and a range of effective instructional strategies. Given this set of beliefs about value of the standards, their oft-stated admonition NOT to use the standards as a checklist tool for summative beginning teacher evaluation should come as no surprise.

Final Revisions with the Validity Study

Based on specific suggestions from the focus groups, many indicator questions were deleted, revised or reorganized within and across domains. These revisions attempted to clarify the language used to describe specific aspects of teaching practice and to communicate relationships within and across domains. The stems of the indicator questions were also changed to give the questions greater depth and to allow teachers to think about them from different angles.

The final draft of the Draft Framework revision, named by the authors *The California Standards for the Teaching Profession*, was completed at the end of August, 1996, and submitted to the BTSA Interagency Task Force for review and editing. Two members of the Task Force returned recommendations at that time and final changes were made.

We also received a set of recommendations for revision from the SB 1422 Advisory Panel. The 1422 Panel and its subcommittees represented the concerns of various constituents in the areas of computer education, critical thinking, gender equity, health, mainstreaming, parent involvement, reading, school violence, and self-esteem. While these recommendations did not emanate from the Framework

Restoring Balance

validity study, the panel felt they merited inclusion because they helped to broaden or strengthen the definition of teaching practice within the contexts of California's changing demographics and the needs of the greater community.

Implications of the Validity Study

The results of the first phase of the validity study revealed implications for BTSA programs and preservice education. Three main concerns surfaced as the team analyzed the data and revised the document:

- u inconsistencies in teachers' and support providers' conceptual understanding of the underlying constructs and terminology in the document;
- u the relevance of the Standards to the bigger picture of support and assessment, including issues surrounding how the Standards would be used;
- u implications for preservice programs and for initial teacher certification.

Of these concerns about the Standards, the second and third have, perhaps, the greatest impact on current developments in California. Issues of teacher education, induction, and teacher evaluation are immediate, and all involve the CSTP. The concerns about the content, construct and consequential validity of the Standards voiced in the 1995-96 validity study have enormous implications now that the CSTP serve as the foundation for most major reform initiatives in teacher preparation and professional development.

Sensing the potential impact of the CSTP, the validity study recommended several "next steps." Consistent with the Task Force's recommendation, the validity study team recommended the creation of a comprehensive "package" of support and resource materials to accompany the Standards. Foremost among the supporting materials were to be developmental scales, a set of critical teaching incidents or vignettes to illustrate aspects of best practice, a guide to appropriate usage of the Standards in support and assessment, and a glossary of terminology. The purpose of the package would be to provide educators with a range of tools for promoting reflective practice, the support and assessment of beginning teachers, and professional development. Thus, the research team argued, the package would play an essential role in the constructive enactment of the Standards.

Phase Two of the Validity Study—Late 1996-1997

The initial scope of work for the validity study included a second phase that called for the crafting of developmental scales and a validity study of those scales. As this work began to unfold, the state legislature acknowledged the value of BTSA to beginning teachers by granting over one hundred million dollars to BTSA for fiscal year 1997-98. Plans for the expansion of BTSA into a statewide program were underway with reorganization to create regional clusters of BTSA programs. Increased emphasis on the role of assessment and the need for stronger accountabil-

ity accompanied legislative funding. At this time, Interagency BTSA Task Force meetings turned more frequently to discussion of the need for a comprehensive, coherent, assessment system that would standardize the assessment of beginning teachers in BTSA programs and make each site more accountable for its process. The debate between the summative or formative nature of BTSA assessment again found its way tensely to the table. References to a high-stakes component to the assessment model emerged more frequently as a theme. Several members of the Task Force forcefully suggested the development of summative benchmark evaluations for teachers at the end of the second year of teaching.

Meanwhile, the Phase II Validity Study research team continued its work drafting developmental scales to accompany the new standards. These scales would define 3-5 levels of beginning teacher performance at the sub-domain level. The research team had been charged with drafting developmental indicators and descriptors for each subdomain. The scales were drafted and aligned to the CSTP by two members of the research team, then reviewed by the full research team and members of the Interagency Task Force. The authors presented the scales to two structured focus groups for feedback where they were received with enthusiasm. A full validity study on the scales, as originally outlined in the 1995 Augmentation Grant Scope of Work, was never conducted. The data from the focus groups were analyzed and used in the revision of the draft. The Draft Developmental Scales and a Final Report on their development were submitted to the Task Force in July, 1997. Consistent with all previous work on the standards and scales, the Final Report recommended including an introductory letter, (deemed critically important to define the purpose and provide context for the scales to users); a glossary of terms, and practical descriptions or vignettes of teaching to illustrate the practice represented in the scales.

In addition, the Final Report recommended continued research through a validity study that would follow the CSTP and the scales into classrooms to document their validity and use. The authors noted that "our current work established both a theoretical construct for the scales as well as focused feedback from multiple perspectives on the content validity of the scales, (but) it did not systematically observe teachers over time and then assess the degree to which the scales match the actual development of teachers" (Final Report on the Developmental Scales, July, 1997, p.9).

The other major recommendation in that Final Report emphasized the consequential validity of the scales:

In focus groups on the Standards and scales over the past two years, teachers and teacher educators expressed significant concern over the potential uses and abuses of these tools. In order to have some assurance that the Standards and the scales constructively influence opportunities for professional development (support and assessment) for teachers, and subsequent learning for students, they will require rigorous inquiry ... to determine 1) their use in guiding professional development,

Restoring Balance

support and assessment, and 2) the conditions that support and constrain their constructive use. (Ibid, p.9)

Changes to the CSTP

In the summer of 1997, the State Board of Education endorsed the *California Standards for the Teaching Profession*. This marked the completion of the adoption process and official publication of the document followed. Since its official adoption, several changes have been made in the document that reflect their changing uses and the changing players. Later versions of the CSTP, for instance, do not include the footer defining references to diversity and inclusion that accompanied the original document. In addition, the organization of the Introduction is changed with the emphasis on the holistic and developmental view of teaching moved from the beginning of the Introduction to the end. Another component, titled “Diversity of Teaching in California,” is subsumed within the closing section. Although these revisions may suggest a particular logic, they also shift the weight of the original intent of the Introduction which, in large part, reflected the thinking of those who developed the Standards and the findings of the validity studies.

The CSTP After the Validity Study—1997 to Present

While phase two of the validity study continued, other events took place. In December of 1996, members of the SCNTP and the research team met to prepare for the last Task Force meeting of the year. At this meeting, the CCTC presented their proposal for a comprehensive assessment model, called then, “The Ideal Assessment System.” This proposal was clearly established as the operative plan for the state, although the SCNTP also submitted a plan for an holistic, integrated support and assessment system based on the CSTP.

The emphasis in the discussion following the CCTC’s presentation was on BTSA’s expanding role and how the 1422 Panel might recommend a summative assessment at the end of the first semester of a teacher’s second year. The group stressed the need to coordinate all components of the support and assessment process and to make the training of support providers a central focus to ensure their ability to work knowledgeably with the new assessment system. The group was also informed that ETS had agreed to adapt their Pathwise assessment system to meet California’s needs, opening the possibility of incorporating Pathwise into the new system. This presented a counterpoint to the role of the FWL (now WestEd) portfolio, funded by the state and being developed as the central component of the BTSA assesement system. This could be said to mark the initiation of a new phase of assessment design and development which would become the California Formative Assessment and Support System for Teachers (CFASST).

The work of the Phase II validity study came to an end in summer, 1997, with the submission of the scales to the BTSA Interagency Task Force, and work then

began on CFASST. In September, 1997, the CFASST Year One Design Team was established to design and develop a comprehensive assessment system. The Year One Design Team was composed of BTSA Directors, some of whose programs used local assessment systems and some whose programs used ETS' Pathwise, and representatives of the CCTC, CDE and finally, ETS, who brought to bear their significant financial and technical resources regarding psychometrically sound and legally defensible assessments, and to a large degree, determined the timeline for production of CFASST.

By February, 1998, the Design Team began revision of the CSTP-based scales originally written as part of the 1995-97 validity study. The final revision of the scales, reconstructed in CFASST as the Descriptions of Practice (DOP), was strongly influenced by ETS during the development of the first-year CFASST cycle. Concerns about the content, language, and consequential validity of the scales, echoing concerns for the uses of the CSTP, continued as a recurring theme throughout the early development of CFASST, particularly in regard to the scales.

As the CSTP make their way through the induction and professional development system, they continue to be presented in contexts and formats that differ from those originally developed. Since 1997, there have been versions of the CSTP and its graphic organizers published in formats antithetical to the findings of the 1995-96 validity study. In one notable instance, the CSTP is presented as a central element of the California Department of Education's professional development document, *Designs for Learning*, published in 2000. *Designs for Learning* serves as a "resource to help teachers develop and implement a professional development plan that will significantly affect what they do in the classroom." There is an overview that describes the CSTP as being based on current research. Following this introduction is a matrix of six boxes, each containing one standard, and each labeled "Standard One," Standard Two," etc. The indicators in each Standard are numerically coded. This matrix was copyrighted by the CCTC and the CDE in 1998. This change in format stands in stark contrast to the findings of the validity study which stated:

In the original (format), the domains appeared in numerical sequence suggesting that they were to be addressed by teachers in that particular order and implying that teaching is a set of sequenced, discrete tasks. The data from the focus groups strongly suggested that the numbers be removed so that the document reflects a more integrated, holistic view of teaching.

In addition, as previously noted, the Draft Framework was accompanied by a one-page summary matrix of the six domains and their corresponding subdomains. Designed as a grid of six squares, this matrix reinforced the notion of discrete and sequential tasks. The focus group participants suggested a revised graphic representation that reflected a more holistic view. The validity study team developed several graphics to support this view. (Framework Validity Study Final Report, 1996, p.45)

Restoring Balance

The matrix or “placemat” presented in *Designs for Learning* resembles the old Draft Framework grid and a matrix used by the ETS Praxis and Pathwise assessment systems far more than it resembles what was called for in the validity study or produced in response to it. This, perhaps, reflects the role ETS played in the development of CFASST (1997-99,) the beginning teacher assessment system, of which the CSTP is the foundation.

While the text of the Standards has not so far been changed, the State has published revised versions of primary supporting documents, as representations of the content and construct of the CSTP, in forms that contradict the findings of the research upon which the Standards are founded.

Conclusions

Teaching policy is increasingly seen and used as a lever for reform of public education in the state of California and across the nation. Public education in a democratic society requires a balance of multiple perspectives and interests. Those required perspectives include all branches of the government (e.g., the administrative, the legislative, and the judicial), the profession, the parents and families of children, the business community, and, perhaps most of all, the strengths, interests, and needs of children. The evolution of the *California Standards for the Teaching Profession* demonstrates how shifting the relative balance of perspectives and interests at play in the making of teaching policy can influence the nature and uses of the resulting policies.

From the beginning of the Standards’ development, state-level policy makers were pursuing a vision of support for all beginning teachers which would, in their way of thinking, require both greater uniformity across contexts as well as greater accountability for results. Accountability means different things to politicians than it does to teachers and teachers of teachers. To politicians, without contextualized knowledge, and relying on a single, “snap shot” measure of success, it often means numbers, preferably a “bottom line” number that proves their policy worked. The business community holds an analogous definition: numbers prove that money was well spent and the “products” of the system can do the kind of work that makes money for the company. To teachers, immersed in contextualized knowledge and with a longer frame of time, it means ALL they see in a year or more regarding each and every individual in their care.

In addition, the courts influenced the process. As soon as stakes are attached to assessments of standards, legal action follows. The courts require that educators follow prescribed rules. In the case of assessments, those prescribed rules have been developed by psychometricians for whom teachers’ knowledge of their students is bias and contextual differences must be, as much as possible, eliminated in the interest of comparability of scores. Embroiled in an extraordinarily expensive defense of CBEST, their basic skills test, throughout the development of the CSTP,

the CCTC was acutely aware of these rules and the implications if they were not followed. Each of these definitions of accountability is accompanied by a set of assumptions about the nature of teaching and learning. Not surprisingly, the more contextualized the perspective, the more teaching is conceived of as complex, holistic, interrelated and learning conceived as a complex set of developmental processes. The more decontextualized the perspective, the more teaching is conceived as atomistic and divisible into discrete behaviors and learning conceived of mechanistically and linearly.

During the evolution of the CSTP, these perspectives were differentially represented. (Interestingly, parents and students seemed to have had a nearly negligible role in the development of the Standards, and the policies and practices that have been developed around them.) Originally embedded in the work of a small cadre of beginning teacher support providers and researchers holding closer to an educator perspective than a psychometric perspective, the Standards reflected an educational practitioner's focus. The purpose of the Standards was to support beginning teachers. The best way to do so was to let support providers use what they knew in the way they knew best. The work was not about evaluation, it was about support, and in some ways, definition limited support. When researchers and the legislature pushed for tighter definition for what that meant, the result was the initial set of frameworks. There was little clout to them, and, for the most part, the BTSA community used them to support their efforts.

The 1994-95 Task Force appointed by the Department of Education and the CCTC consisted of a broader array of the educational profession than the original CNTP community. This group was more willing to codify the Standards and see them used to shape teaching policy. Still, they were mostly educators who perceived teaching and learning in ways consistent with, if not identical to the CNTP community. They were willing to risk the potential for abuse, working on the assumption that the profession had a responsibility to govern itself and the first act of professional oversight would be a set of standards developed by the profession. They still perceived, however, the danger of standards without opportunities for learning and workplace conditions that allowed educators to meet professional standards. Thus, this group focused on using the standards to "educate the profession" (their unwavering devotion to a "preamble" or introduction) and to "educate policy makers" (their unwavering commitment to never let the standards see the light of day without the accompanying materials).

The validity study team kept working primarily within the profession. In some ways, the findings of the validity study remained consistent with the previous work. The systemic inquiry conducted by the research team continued to support an educational perspective (e.g., contextualized, holistic, developmental, supportive) versus a political, psychometric, or business perspective (e.g., decontextualized, atomistic, mechanistic, evaluative). In fact, the validity study results are closer to the tone of the original work than the work of Bartell's 1994-95 Task Force. Its

Restoring Balance

findings suggest that practicing educators might be less supportive of a profession assuming the responsibility for governing itself than were the “educational leaders” appointed to the Task Force. Ultimately, however, all groups who had a hand in developing the CSTP through the validity study had similar professional perspectives on teaching and learning, and expressed profound fear if these standards were to be used in ways incompatible with the conceptions of teaching and learning which undergirded them. Finally, the validity study completed by educational researchers (not psychometricians) expressed deep concern that the Standards, in the form of developmental scales, were about to be used with no empirical basis whatsoever.

The subtle shifts in balance occurring between the work of the first three stages of the evolution of the CSTP pales in comparison to the shift between the validity study and what occurred thereafter. In short, the political and psychometric perspectives shifted to the fore and the professional to the rear. The representation and uses of the Standards since the completion of the validity study have reframed their underlying conceptions away from their original positions and born out some of the fears expressed earlier in their evolution.

The purpose of this discussion is not to argue that politicians, business people, and psychometricians have destroyed the CSTP. An equally valid argument can be made that the profession abdicated its responsibility by refusing to evaluate its members or at least monitor entry into the profession. From this perspective, the State is upholding its legal and ethical responsibility to the children it compels to school each day. Our purpose is to have all of the parties who should and must be involved in creating and enacting teaching policy, work together and understand each other in the support of students. We all have a role in restoring the healthy balance public education, and all our children, require. Psychometric rules must at times be reinvented in the interest of children. Simple and quick political solutions must at times be tempered in the interest of children. The tendency to impose mechanistic models on the complex and holistic process of teaching and learning must cease in the interest of children. Finally, as educators we must think through our own practice and our own involvement in policy at all levels of the educational system. We must take hard looks at our own responsibilities towards the profession in the interest of children.

References

- Agenda of the meeting of the California Commission on Teacher Credentialing. January 1997. Sacramento, CA: Commission on Teacher Credentialing
- California Department of Education/California Commission on Teacher Credentialing (1996) *California Standards for the Teaching Profession*. Sacramento, CA: California Department of Education.
- California Commission on Teacher Credentialing /California Department of Education. Augmentation Grant for Development, Implementation, and Support for the California Teacher Support and Assessment Program. March, 1995.

- Charge to the Task Force: Development of the Framework of Knowledge, Skills, and Abilities for Beginning Teachers in California." California Commission on Teacher Credentialing and California Department of Education. March, 1994.
- CSTP numerically coded matrix - I. Adapted from "California Standards for the Teaching Profession," pp 5-22. California Commission on Teacher Credentialing and California Department of Education. 1997.
- CSTP numerically coded matrix - II. "California Standards for the Teaching Profession." California Commission on Teacher Credentialing and California Department of Education. 1998.
- "Descriptions of Practice" (DOP), "California Formative Assessment and Support System for Teachers." California Commission on Teacher Credentialing and California Department of Education. May 1998.
- "Designs for Learning: An Introduction to High Quality Learning." California Department of Education. 2000.
- "Draft Framework of Knowledge, Skills, and Abilities for Beginning Teachers in California." Commission on Teacher Credentialing and California Department of Education. August 1995.
- Educational Testing Service. 1993. "The Praxis Series: Professional Assessments for Beginning Teachers." Princeton, NJ: Educational Testing Service.
- Garmston, S. (1992) "Success for Beginning Teachers", p.4 cited in "Charge to the Task Force: Development of the Framework of Knowledge, Skills, and Abilities for Beginning Teachers in California." Commission on Teacher Credentialing and California Department of Education. March 1994.
- Moir, E., & Garmston, S. (1992) Continuum of Skills, Knowledge, and Attitudes. "Continuum of Skills, Knowledge and Attitudes." Adapted from work developed by Dr. Kendyll Stansbury, Far West Laboratory; Dr. Betty Ward, Southwest Regional Laboratory; Dr. Mary Diez, Alverno College. Santa Cruz, CA: Santa Cruz New Teacher Project.
- Moir, E., Snyder, J., Freeman, S., & Whittaker, A. (1996) Final report on the Phase I Validity Study of the Draft Framework of Expectations for Beginning Teachers. Report to the BTSA Joint Task Force as a part of the BTSA Augmentation Grant awarded to UC Santa Cruz BTSA for Development, Implementation, and Support for the California Teacher Support and Assessment program.