A Coherent System of Teacher Evaluation for Quality Teaching

Accomplished California Teachers:
Tammie Adams, Elena Aguilar, Ellen Berg, Liane Cismowski, Anthony Cody,
David B. Cohen, Sandra Dean, Lynn Formigli, Jane Fung, Cliff Lee,
Kathie Marshall, Nancy Skorko, and Silver White
Stanford University
United States

Abstract: This article addresses the questions, "How should we evaluate the quality of teaching?" and "What kind of evaluation system will move all California teachers on a path of improvement throughout their careers?" The article, adapted from a report written by a group of accomplished California teachers, recommends seven core principles to develop a system of teacher evaluation that provides meaningful and ongoing inputs and moves teachers on a path of continuous improvement throughout their careers. It also addresses some of the elements of accountability that the state system writ large can adopt to create a system that fosters high-quality teaching.
Keywords: teacher evaluation; quality; accountability

Un Sistema Coherente de Evaluación Docente Para una Enseñanza de Calidad

–

Un sistema coherente de evaluación docente para una enseñanza de calidad
Resumen: Este artículo aborda las preguntas, "¿Cómo debemos evaluar la calidad de la enseñanza?"
y "¿Qué tipo de sistema de evaluación movilizaria a todos los docentes de California hacia un
camino de mejorias a a lo largo de sus carreras?" Este artículo, adaptado de un informe escrito por
un grupo de docentes de excelente trayectoria en California, recomienda siete principios básicos para
desarrollar un sistema de evaluación docente que proporciona recursos significativas y promueve
una trayectoria de mejora continua a lo largo de sus carreras. Asimismo este trabajo discute algunos
de los elementos de un modelo de responsabilidad educativa que el sistema estatal puede adoptar
para crear un sistema que fomente una enseñanza de alta calidad.
Palabras clave: evaluación docente; calidad; modelos de responsabilidad educative

Um sistema coerente de avaliação docente para uma educação de qualidade
Resumo: Este artigo aborda as questões, "Como devemos avaliar a qualidade do ensino?" e "Que
tipo de avaliação mobilizaria a todos os professores na Califórnia para um caminho de melhora ao
longo de suas carreiras?" Este artigo, adaptado de um relatório escrito por um grupo de professores
de desempenho excelente na Califórnia, recomenda sete princípios básicos para desenvolver um
sistema de avaliação docente que fornece recursos significativos e promove um caminho de melhoria
contínua ao longo de suas carreiras. Este artigo também discute alguns dos elementos de um modelo
de responsabilidade educacional que o sistema estatal poderia tomar para criar um sistema que
estimule uma educação de alta qualidade.
Palavras-chave: avaliação de professores; qualidade; modelos de responsabilidade educative

Introduction

Evidence indicates that teachers are the most critical in-school factor for student success –
a point on which there is broad consensus across the political spectrum. But how to promote high-
quality teachers is a point of more contention and, for the past 14 years, has been driven largely
through the narrow, high-stakes approaches of the No Child Left Behind act. During this period,
the country has also seen a growing shortage of qualified science and math teachers (Sterling, 2004)
and high-quality teachers in inner city schools.

The predominant strategy of the past decade for improving teaching quality and for bringing
the best teachers to the country's most disadvantaged students has backfired. But there are effective
and coherent approaches that work. In their 2014 paper, Darling-Hammond, Wilhoit, and Pittenger
outline the critical policy points that can set the foundation for a new system of accountability to
address America's widening opportunity gap in education. The systemic aproach, they argue, should
rest on three pillars – accountability for meaningful learning, adequate and wisely used resources,
and professional capacity.

The latter aspect – professional capacity that enables professional accountability – was
addressed comprehensively in 2010 by a team of California-based teacher leaders, who were part of
the Accomplished California Teacher (ACT) network operating under the auspices of the National
Board Resource Center at Stanford University. The ACT members are widely recognized as expert
teachers; most are National Board certified; many have won national and local teaching awards. All
are deeply committed to building a profession of teaching that can serve all students well.

Together, they conceived a system of accountabilty for teaching that can develop highly
skilled teachers and bring the best teachers into the classrooms of the most disadvantaged students.
Following is the summary of their concerns and recommendations, derived from the original report.
While the specific tools and programs mentioned here now have an additional five years of data that

can be examined, the foundational principles are a powerful model that any state can look to to develop a highly qualified teaching force.

The full text of their recommendations, *A Quality Teacher in Every Classroom: Creating a Teacher Evaluation System that Works for California,* can be found at https://edpolicy.stanford.edu/publications/pubs/108.

Why Evaluation Systems Should Change

While evaluation processes across the California vary widely, many of them look very much the same as they did in 1971 when the California Legislature enacted the Stull Act governing teacher evaluation. Among the common chllaneges: a system that teachers do not trust, that rarely offers clear directions for improving practice and that often charges school leaders to implement without preparation or resources.

Jane Fung, National Board Certified Teacher and Milken award winner in the Los Angeles Unified School District, shared her experiences with evaluation as she has experienced them in her career over the last 20 years:

> I have had administrators who never came into my classroom for formal observations or asked me for anything more than the initial planning/goal sheet. I have had administrators observe a formal lesson and put the feedback sheet in my box without ever having spoken to me about the lesson, and I have had years where I am just asked to sign the end-of-the-year evaluation sheet [without being observed].

Ellen Berg, a Middle School teacher in San Diego, shared a different experience – more intrusive but no more productive:

> Because there is not a common language about what quality teaching is, in some cases we use a checklist of random things. In San Diego Unified they had us go visit classrooms with a list of all these things that were supposed to be going on – group-work, cooperative learning, etc. – and it was impossible to do all these things in a 15- (or even 50-) minute period, and teachers were being ripped up for not doing everything on the list.

While there are places where evaluation actually helps teachers improve their practice, such examples are rare in California and across the nation. The costs of the existing systems, both to the fiscal bottom line and to the teaching profession, are large. The fiscal costs entail much more than those associated with removing poor teachers after the tenure deadline passes. The financial impact accrues to school districts that must replace teachers who leave due to dissatisfaction with the profession caused by lack of guidance about improving their work and to loss of leadership that is overwhelmed by the task of providing that guidance. In cases where teachers leave because of lack of guidance and support, the costs related to hiring each new faculty member can amount to upwards of $20,000 (Barnes, Crowe, & Schaefer, 2007). The impact on the profession comes with the loss of potential talent that disappears when promising young teachers leave. It comes as a result of the lack of opportunities for teachers to master the craft of teaching and advance their effectiveness with students, mastery that comes from collaboration with expert evaluators.

If we are to develop an effective system of teacher evaluation, many things need to change, including the following:

- Teachers and their evaluators do not hold in common truly well-defined and detailed pictures of what constitutes good professional practice at each level of teacher development. Clearer, more elaborated standards are needed.

- The focus of evaluations is not on improving the quality of teaching. There is rarely substantive discussion that occurs either before or after an observation that is focused on ways to get better at teaching. In most cases, the evaluations are conducted for compliance, not improvement.
- The amount of time available for principals to conduct effective evaluations is seriously limited, particularly in large schools and in high-need schools where the administrative demands are large. Furthermore, the amount of preparation principals receive in doing evaluations is inadequate. One evaluator in a school is rarely sufficient to judge the skill of teachers across a range of content and developmental levels, no matter how well-resourced a school might be.
- Most evaluations pay little or no attention to the performance of a teacher's students, even though state legislation requires student outcomes be considered. Evaluations too often focus on easy-to-observe practices like classroom management and whether students are on task, rather than looking for evidence that students are actually mastering the learning goals set for them.
- Current evaluation procedures occur on schedules mandated by local agreements that are not considerate of teachers' actual needs and have no sense of urgency about which teachers' work needs more careful support or scrutiny.
- Most evaluations are not used to target the needs of individual teachers and help them select professional development to address those areas in which they need additional knowledge or skills. This further contributes teachers' views that evaluation is not about their developing mastery of professional standards, but is rather a routine designed to ensure that an administrator is performing his job.

An Evaluation System Based on What We Know about Good Teaching

The primary purpose of evaluation must always be the improvement of teaching and promotion of better student learning. This is the overarching principle upon which a new model should be built. Indeed, at the most fundamental level what we want is honest evaluation of our work by skilled and knowledgeable evaluators who can help us see the ways to improve practice at every stage of our professional lives and increase our contributions to the learning of our students. An effective evaluation system will be built on that one overarching principle.

We recommend the following seven principles as the foundation for the coherent, reliable system of teacher evaluation we want and need:

1. Teacher evaluation should be based on professional standards and must be sophisticated enough to assess teaching quality as it is manifested across the continuum of teacher development. The state should use the California Standards for the Teaching Profession and the National Board for Professional Teaching Standards (National Board) standards to create a continuum of expectations from pre-service teaching to accomplished practice. The standards of teaching practice selected as appropriate at each level of teacher development should guide evaluations while accounting for the requirements for successful teaching in the variety of unique contexts in which teaching practice occurs.
2. Teacher evaluation should include performance assessments to guide a path of professional learning throughout a teacher's career. Existing assessments like the Performance Assessment for California's Teachers (PACT) and the Teacher

Performance Assessment (TPA)[1] in pre-service; a similar, more productive tool to guide induction through BTSA; new, authentic assessments related to classroom practice for developing professionals; and the National Board assessment should be used to build coherence and continuity.

3. The design of a new evaluation system should build on successful, innovative practices in current use, such as evaluations built on teachers' self assessments in relation to high standards of performance or evidence-based portfolios that demonstrate ways that a teacher's instructional practice is contributing to student achievement. Teachers must have a significant role in the design of a new framework and in promoting it among teachers in the state.

4. Evaluations should consider teacher practice and performance, as well as an array of student outcomes for teams of teachers as well as individual teachers. Those outcomes should include a number of indicators not limited to standardized test scores, that constitute evidence of student success. Evidence should include performance on authentic tasks that demonstrate learning of content; presentation of packages of evidence from formative assessments that show patterns of student improvement; the development of habits of mind that lead to improved academic success along with contributing indicators like attendance, enrollment in advanced courses, graduation rates, pursuit of higher education, and work place success. Teachers should be evaluated both on their success in their own classroom and their contributions to the success of their peers and the school as a whole.

5. Evaluation should be frequent and conducted by expert evaluators, including teachers who have demonstrated expertise in working with their peers. Evaluators at each juncture should be trained in the recognition and development of teaching quality, understand how to teach in the content area of the evaluated teacher, and know the specific evaluation tools and procedures they are expected to use. There should be training opportunities available for evaluators and final recommendations to teachers should be subject to review by a reliable evaluation oversight team.

6. Evaluation leading to teacher tenure must be more intensive and must include more extensive evidence of quality teaching. This evidence should be collected and reviewed by both the teacher and trained evaluators and should include documentation that shows that the teacher's practice exhibits the standards that define good practice. The process should be an ongoing part of a serious teaching induction process that helps novices grow in their profession, with the help of mentors and coaches, guided by clear standards of practice.

7. Evaluation should be accompanied by useful feedback, connected to professional development opportunities, and reviewed by evaluation teams or an oversight body to ensure fairness, consistency, and reliability.

Principle 1

Teacher evaluations must be based on professional standards that are sophisticated enough to match our knowledge of teaching quality as it is manifested across the whole continuum of teacher development. The state should use the California Standards for the Teaching Profession (defined at a more detailed level to make evidence of them unambiguous) and the National Board standards to create a continuum of expectations from pre-service teaching to accomplished practice. The standards of teaching practice selected as appropriate at each level of teacher development

---

should guide evaluations but should take into account the requirements for successful teaching in the unique context in which it occurs.

> The evaluations themselves are typically of little value – a single, fleeting classroom visit by a principal or other building administrator untrained in evaluation, wielding a checklist of classroom conditions and teacher behaviors that often don't even focus directly on the quality of teacher instruction. It's typically a couple of dozen items on a list: "Is presentably dressed," "Starts on time," "Room is safe," "The lesson occupies students…." But, in most instances, it's nothing more than marking "satisfactory" or "unsatisfactory."
>
> – Michigan State University Professor Mary Kennedy
> (Toch & Rothman, 2008)

As we explored the concept of teacher evaluation, we wrestled with the dilemma of how to articulate our shared understanding of the complexity of teaching and to clearly identify the elements that comprise high quality teaching. We turned to experts in the field first and found some basic premises that resonate with our own experience and views of quality teaching.

In 1988, Lee Shulman described something he called "pedagogical content knowledge," which "transcends mere knowledge of subject matter as well as generic subject matter of pedagogy alone." He explains:

> The teacher not only understands the content to be learned and understands it deeply, but comprehends which aspects of the content are crucial for future understanding of the subject and which are more peripheral and are less likely to impede future learning if not fully grasped. The teacher also understands when present preconceptions, misconceptions or difficulties are likely to inhibit student learning. The teacher also has invented or borrowed or can spontaneously create powerful representations of the ideas to be learned in the form of examples, analogues, metaphors or demonstrations. (p. 37)

Mere understanding of the content is only half the challenge. Teachers must translate that content into language and ideas accessible to their students.

Darling-Hammond (2007) identifies a list of qualities that have been found by research to contribute to teacher effectiveness. These qualities indicate the complexity of teaching and how thoughtful approaches to evaluation must be constructed. The qualities include:

- Strong general intelligence and verbal ability that helps teachers organize and explain ideas, as well as observe and think diagnostically;
- Knowledge of how to teach others in a subject (content pedagogy), in particular how to use hands-on learning techniques (e.g., lab work in science and manipulatives in mathematics) and how to develop higher-order thinking skills;
- An understanding of learners and their learning and development – including how to assess and scaffold learning, how to support students who have learning differences or difficulties, and how to support the learning of language and content for those not proficient in the language of instruction; and
- Adaptive expertise that allows teachers to make judgments about what is likely to work in a given context in response to students' needs.

Obviously, these attributes of high quality teaching are not acquired all at once and do not arrive fully developed on the first day a teacher enters the classroom. A set of qualities this complex demands a thoughtful approach that supports teachers to develop and improve throughout their

careers. If indeed this is our goal for all members of our profession, then we must find ways to reach that goal using many of the same approaches to formative assessment we know work with students. An essential part of that kind of formative assessment, as we pointed out earlier, is frequent feedback from and collaboration with knowledgeable peers that help us see our own teaching clearly in light of high standards of practice and reflect on ways to improve. Many of us have personally experienced holding up our practice against the indicators of quality teaching based on the standards of the National Board, which contain just those indicators described by Shulman and Darling-Hammond. A good evaluation system will build such experiences into a continuum of teacher development for all California teachers.

Principle 2

Teacher evaluations should include performance assessments to guide a path of professional learning throughout the career. Existing assessments like the TPA and PACT in pre-service, a revised, more productive tool for BTSA, and the National Board assessment should be used to build coherence and continuity.

> Many times since (undertaking National Board certification), I have thought about these two polar experiences. I feel they shine a light on key inadequacies of the current Stull evaluation process that might be thoughtfully improved. In order to be effectively implemented, the Stull process needs to reflect meaningfully the performance of teachers at every level – from the weakest, who must be identified for assistance or dismissal – to the superior, who deserve recognition for their exemplary roles in our schools. The media frequently clamor about getting rid of "bad teachers," which taints us all. Meanwhile, many of our most outstanding teachers work tirelessly year in and year out with little official recognition or reward for their efforts.
>
> – Jane Fung, National Board Certified Teacher,
> Los Angeles Unified School District

We turned our attention next to what a system based on solid indicators of teaching quality might look like and how, if it were used well in the hands of knowledgeable assessors, it would help teachers assess their strengths and needs, set goals for improvement, choose instructional strategies, work with colleagues to build new knowledge, self-assess their own work, reflect on the quality of the work by looking at student success, and begin the cycle anew in the quest for continual improvement.

Danielson (1996) describes a framework that organizes teaching standards into four domains:

- planning and preparation
- classroom environment
- instruction
- professional responsibilities

She provides rubrics for observing and reflecting on practice in each of these domains. Danielson (1996) explains, "Teaching is so complex and its various components so intertwined that many novices feel overwhelmed. A framework for teaching offers a structure to assess a teacher's practice and to organize improvement efforts." We suggest that veterans often feel overwhelmed, as well. The complexities of teaching do not diminish as we learn more. Indeed, one National Board Certified Teacher recently said that one of the results of the certification process that surprised her

was that, as she began to think more deeply about her practice, the number of questions she raised about her work grew exponentially.

Frameworks for observing and assessing teaching are already used in many places and circumstances in California. The PACT assessment uses a framework appropriate for the end of pre-service study. BTSA, in many of its forms, uses a framework that is based on the California Standards for the Teaching Profession. The National Board uses a well-structured framework adapted to a variety of student developmental levels and content areas and has a rigorous process of peer assessment that is well respected for it reliability as a measure of teacher accomplishment.

Evaluation within school districts would be improved if these standards were used. The glaring hole we, like many others who have looked at the issues, would point out is that the majority of teachers in the state are not evaluated in a reliable and consistent way using any framework based on standards for the profession. In the same way that good teachers help students understand learning goals that include detailed descriptions of the expected performance accompanied by exemplars of that performance, teachers should be provided with the same clear expectations in the form of elaborated descriptions of standards, exemplars of good practice, a framework for evaluating their work, and a process for feedback from other knowledgeable professionals. If the state is to be taken seriously about achieving the goal of uniformly high quality teaching at all levels and content areas, then finding a way to evaluate teachers using frameworks aligned with standards like those already created and in use at the foundational levels is essential. There must also be ample support for teachers at every career stage as they work to meet and exceed those standards.

In our discussions about the ways that standards are used and understood in our various workplaces, there was agreement that most schools in the state know about the standards, but there is little agreement that they are a force in the work of teachers in classrooms or the reference points that drive conversations about teaching practice. As Danielson (1996) states:

> With a framework of professional practice in hand… participants can conduct conversations about where to focus improvement efforts within the context of shared definitions and values. These conversations can focus on means, not ends, and they can be conducted in an environment of professional respect.

One key element that has been left out of the state's California Standards for the Teaching Profession is in the "making sense" aspect of the standards. Our experiences in working toward National Board certification included extensive time devoted to making sense of the standards and identifying evidence of them in our own practice. That process is what National Board Certified Teachers credit for transforming their practice. As teachers who have experienced this transformation, we are eager to see the same kinds of conversations included in evaluations across the professional development continuum because we know that when communities of teachers engage in them together, the quality of teaching improves. Danielson (1996) states this well:

> When teachers engage in self-assessment, reflection on practice, and professional conversation, they become more thoughtful and analytic about their work, and are in a position to improve their teaching. Evaluators can contribute to teachers' professional learning through the use of in-depth reflective questions. By shifting the focus of evaluation from "inspection" to "collaborative reflection" educators can ensure the maximum benefit from the evaluation activities.

The challenge is to do what has been done at the pre-service level in the design of PACT and TPA and create performance assessments that span every level of the continuum including National Board certification and beyond. A framework based on common standards must be adaptable for use in the evaluation of teaching in many different contexts, content areas, and developmental levels. We believe that the California Standards for the Teaching Profession can fulfill their potential to be that driver of higher teaching quality in the state.

Principle 3

The design of a new evaluation system should include a study of existing successful innovative practices that are being used in the state and country and should encourage schools and districts to explore innovations. Teachers must have a significant role in the design of a new framework and in promoting it among teachers in the state.

> Evaluation [now] is based on a culture of fear – a "gotcha" mentality. Teachers are afraid of being observed by peers, and afraid of anybody being critical, and then that observation of their failures going into their regular evaluation file.
> – Tammie Adams, Oakland Unified School District

There are examples of innovation within the state that are putting into practice some of the elements of what we have concluded should be part of an effective system of teacher evaluation. The Governor's Committee on Education Excellence, which issued its findings in the 2007 *Students First* report, encourages the idea of learning from local innovations, a few of which we highlight below. These are also noteworthy as examples that contradict the general perception that teacher unions are indefatigable opponents of innovations or reform in teacher evaluation.

In the Palo Alto Unified School District, the local teachers' association and district administration addressed the issue of ensuring consideration of content area expertise in evaluation. There, longstanding practice is in place at secondary schools where teachers serve as instructional supervisors to conduct most evaluations and to take the lead in hiring recommendations. This distribution of the evaluative workload mitigates potential overload for principals. With reference to instructional evaluations in particular, San Mateo Superintendent Scott Laurence, a former principal in Palo Alto, notes the advantages:

> The content area expertise is the most important aspect of teachers evaluating teachers. As a former social studies teacher, I felt comfortable as a principal observing social studies and English classes, and the early-stage courses of math and science. But in upper level math and science courses I had to spend more of my time watching instruction and management, and give less focus to the content. I also believe that curriculum and instructional practices have changed. I was last in a classroom over 10 years ago. It does make a difference.

Santa Clara Unified School District addressed concerns about differentiating evaluations based on the location of teachers on the professional growth continuum. Teachers who have received satisfactory evaluations in the past are allowed to choose the Alternative Professional Growth Evaluation. In this model, teachers choose to focus on improving one aspect of their practice, set concrete goals, and at the end report on what they have learned and how their practice has changed. This allows teachers to engage in what research has shown to be one of the most powerful forms of professional growth available: teacher action research.

Vaughn Next Century Learning Center, a conversion charter school in Los Angeles Unified School District (LAUSD), uses a modified evaluation system somewhat based on Danielson's (1996) framework. It was developed alongside of its performance pay system. It uses a teacher's self-assessment, peer evaluation, and administrator evaluation. It has been an evolving system subject to revisions based on feedback from administrators and teachers together. The approach is particularly interesting not only because of the way that it has evolved over time but also because in its latest iteration it has considered exactly those issues that our group identified as important earlier: consideration of content areas, developmental levels, and contexts. The focus on including evaluation of instructional practices aligned to achieving school-wide targets is a consideration used

at Vaughn that we think deserves careful examination and will be an issue we hope to pursue as we move on to pay-for-performance issues in a future report.

Finally, we considered the innovative system of teacher development used in Poway Unified School District, located near San Diego. Poway has made use of the Peer Assistance and Review (PAR) program to build their approach. Recognizing that some teachers may, at different times in their careers, not be as effective as they should be, California has funded the PAR program for several years. District leaders in partnership with union representatives across the state designed PAR programs to provide a form of assistance and due process for teachers who have received poor evaluations. In most of these programs, a joint committee composed of teachers and administrators oversees a team of coaches (usually classroom teachers on special assignment) who offer assistance to struggling teachers and prepare reports based on their observations. This information can then be factored into decisions made regarding the teacher's professional future. The role of the PAR program in Poway has been expanded to include coaching and reviewing the performance of beginning teachers. Their program, modeled after similar ones in Rochester, New York, and Toledo, Ohio, is in its 21st year.

Charlotte Kutzner, Program Coordinator for the Poway Professional Assistance Program (PPAP), explained the process to our group:

> Poway's PPAP is a BTSA program, and like others, we are responsible for meeting all of the induction standards. But unlike most other induction programs, we are also responsible for evaluation of first-year teachers. So we observe and conference, we support teachers, but our evaluations are not confidential. When I work with a new teacher I share what I see in their classroom with their principal, and also report to a governance board, which includes the assistant superintendent of personnel, the union president, and two teachers. People ask, "How can you do both support and evaluation?" We do, and it has worked since 1987, and from the get-go it has been evaluative. I would say that over the years 95% of the teachers I have worked with, by Thanksgiving, have forgotten that I am their evaluator. I am Charlotte, I am their friend, their colleague; I am there to support them. There is also a program to support veteran teachers who have been rated as not meeting standards. This program, the Permanent Teacher Intervention Program (PTIP) [similar to PAR in other districts], is designed to assist permanent teachers who have been identified as being in serious professional jeopardy. The PTIP teacher receives assistance from a teacher consultant much like the new teacher does in the induction program. In this program, the principal remains the evaluator and the teacher consultant reports progress to the principal and the PPAP Governance Board. Our program is successful because of our working relationship with the district and the union. This is truly a joint effort.

Poway's approach has much to recommend it in terms of efficiency, coherence in evaluation, support at different levels, and building professional relationships, but one of the features that is particularly relevant to the problem of the costs of teacher turnover is the support provided to the struggling teacher. While some have suggested that it is a worthwhile goal to weed out the lowest performing members of the teaching force, the truth is such an approach would be far more costly than a program that intervenes to help low performers rise to levels of proficiency.

At Poway, we see a system in which the teacher evaluator also serves as a mentor. Teachers in Poway have arrived at their model by reasoning that the mentor who works most closely with the teacher is best positioned to make recommendations about the teacher's employment status. The process is open and transparent, and the mentor does not exercise any authority independently of the review board.

Other models stress a separation between the functions of mentorship and evaluation. The reasoning suggests that teachers in need of mentorship must be able to confide in their mentors and risk exposing areas of weakness in their practice as the fear factor takes center stage again. Knowing that the mentor will eventually make an employment recommendation could cause the evaluated teachers to withhold important information compromising the quality of the mentorship and the potential for growth.

Jane Fung, a National Board Certified Teacher in LAUSD, advocates for separation of mentor and evaluator roles from her own experience on both sides of the relationship:

> My first weeks of teaching were full of stress and the feeling of being overwhelmed. I had difficulty with classroom management and wasn't sure how to organize my day, let alone teach the programs I was not familiar with. Luckily, I had a mentor that came into my class and modeled different strategies that she used. She then provided me with a much needed sub day for me to shadow her in her classroom. I was able to gain ideas and the support I needed from her to be a more successful first year teacher. Her role was made clear to me: "I am here to support you, not evaluate you."
>
> When I mentor new teachers, I keep those feelings of my first year close at hand. The idea of being both a mentor and an evaluator for a new teacher feels uneasy to me. As a mentor, I must first establish a relationship and build trust with the new teachers I work with. Confidentiality is essential to building that trust. They must feel safe enough with me to be honest and open about what is going on in their classrooms. As a mentor, I observe the new teacher and document evidence I see, but I share the data collected only with the new teacher, and help them use it to reflect on their practice. Together we develop an action plan on how to improve their practice so that they can feel successful. My role as a mentor is to provide new teachers with support, resources, and a safe place to express their thoughts and feelings. I am a coach, a teacher, and friend to that new teacher. I do not judge, evaluate, or make decisions on their professional career. If I were seen both as evaluator and mentor, I am not sure if I would be able to establish the same kind of trust and provide the support needed to that new teacher; the person meant to support them shouldn't add more stress by evaluating them at the same time.

We are not advocating for any one specific model, but we have agreed that they all have interesting features to recommend them and that getting a system of evaluation right means being willing to look at models being used around the state, and inviting a closer look at them, so that the best ideas can be put together to create the coherent system we need to ensure uniformly high quality teaching. In general the features that they share and which merit inclusion in good evaluation systems are:

- Frequent, ongoing evaluation for new teachers.
- A well-trained evaluator that has expertise in the content areas specific to the teacher's practice and who works collaboratively with other experts to ensure teachers receive accurate and effective evaluation and recommendations.
- The use of multiple measures to evaluate effective practice connected to a wide array of evidence of student outcomes, not just test scores.
- Collaboration between the evaluator and teacher, with a focus on teacher development.
- Differentiated evaluation for experienced teachers, with focus and form open to some negotiation among teachers, administrators, and evaluators (when other than administrators).

We encourage policymakers who will be charged with selecting those who will design a new evaluation system to recognize the expertise of accomplished teachers. We have been involved with change initiatives in various forms at many schools sites. We know from this experience that teachers who do not share some power over decisions made about their work will resort to the power of resistance. Getting the evaluation system right is too important to risk in this way. Therefore, among the principles we believe need to be included in designing a new system is the involvement of teachers at every phase in the process: designing, negotiating, and implementing evaluation. The staff at Vaughn Learning Center echoed this notion in one of their recommendations stating, "Encourage involvement of representative people in designing the system, but recognize the need for a 'change champion' at times."

        We sadly acknowledge that trust between teachers and policymakers has eroded over the last several years as the finger of blame for the collective failures of our schools has pointed at teachers most directly. Whether or not there is justification for this, we believe that renewing the essential trust needed to reform schools is essential. For teachers and their leadership to embrace an evaluation system, school communities and policy leaders must take steps to help repair the climate for our collective learning. In many recent reform efforts, our professional expertise has been denigrated, and professional development has been relegated back to "workshop" approaches designed to ensure compliance with one size fits all mandates.

Principle 4

Schools and districts should focus on building teacher accountability that looks at student outcomes among teams of teachers as well as that of individual teachers. Those outcomes must include agreed-upon indicators, and not only standardized test scores, that are recognized as evidence of student mastery of the state standards for the grade and content area.

        The genuine accountability that we feel to our students and to one another, when we work as part of a functional collaborative community, dwarfs any sense of accountability that can be imposed by test scores, site administrators, or state oversight. This accountability is derived from our shared commitment to the learning and well-being of our students, and our desire to support one another in meeting students' needs. An overarching purpose of evaluation ought to be to promote collaborative examination, analysis, and reflection on the work being done, and figuring out how to improve each teacher's skill set in ways that improve the work of all teachers in the school. As Santa Clara science teacher Lynne Formigli said, "I believe an evaluation system should have at its core the purpose of helping all of us to grow in our profession." Many voice concerns about teachers protecting poor teaching in such an arrangement. On the contrary, we have a very big stake in ensuring that ineffective and poor teachers are either helped to improve in a timely manner or counseled out of the profession. As Lynne Formigli, a National Board Certified Teacher in the Santa Clara Unified School District, said:

        Let's face it! We are much harder on and more demanding of our colleagues than
        any outside evaluator or administrator could ever be. After all, when their work is
        good we all have it easier. When it's not, we all suffer and have much fixing to do.

Student outcomes have the central place in this process. Teachers and administrators should include agreed-upon indicators of learning that are valid and appropriate measures of the curriculum and the students being taught in the evaluation process. There are many instances we know of where this occurs now. Consider the following work undertaken by members of our network:

-   Site-based research is done where teams of teachers collaborate to create common
    assessments focused on important learning goals, and then analyze student learning
    outcomes. They then share instructional strategies used by members of the group that

have worked and analyze the related student outcomes, helping all teachers in the research cohort, if not the whole school, to improve.

- A lesson study project is conducted that involves a team of teachers setting instructional goals based on local needs, observing a colleague deliver a lesson, analyzing resulting student work, and then recommending changes to the teacher's instruction based on student results.
- A collaborative project is designed where teachers of different subjects team up to create an interdisciplinary unit of study, mutually teaching and reinforcing clear learning goals, and then assessing the impact of that approach on student learning and what they might change individually and collectively in their instructional approaches to improve the outcomes.
- A professional learning community functions at the school where teachers meet regularly to review student work in light of current research, best practices, learning objectives and state standards, with a process for sharing findings with colleagues so that they can integrate these ideas in their own practices and improve their instruction.

The kind of collaborative, reflective approach to reviewing curriculum and pedagogy described by Cliff Lee (see sidebar at the end of this section) must occur as an integral part of the work of the school. It is professional development. A consistent complaint about professional development from teachers is that it has been disconnected from real work at school sites. It is as if we all got prescribed the same medicine for whatever disease we might be suffering from. Teachers want to address their shortcomings as long as the solution is connected to the problem and is undertaken in a climate that is based not on fear of uncovering problems but based on getting better at the work they do.

So the question that naturally arises is how does standardized test data fit into teacher evaluation and what dimension of effective teaching does it reveal? It is no secret that many teachers react with skepticism – and yes, fear – to the idea of including student outcomes as part of teacher evaluations or compensation schemes because over the past decade those outcomes have mostly been equated with standardized test scores. This has led to well-documented problems including a narrowing of the curriculum to focus on tested subjects, topics, and even test formats. The more pressure there is to increase scores, the more likely it is that schools will resort to isolated test preparation, which can result in an artificial inflation of scores, while depriving students of the real skills and knowledge they need.

Many of us have had experience with students who have, according to their standardized test scores, mastered the grade- or course-level materials that are prerequisites to what we teach. However, when we begin working with that student on the assumption she has the knowledge base to master the material in our class, we discover that the test score does not match the student's ability to perform at the expected level. What went wrong, we wonder? Many possibilities might be responsible. Among them: the student learned the material only at a level to be able to answer questions on a multiple choice test; there was some guessing or dumb luck involved; the student never had an opportunity to apply the learning in real contexts so it was not retained. Furthermore, standardized tests are frequently invalid for students with special education needs and for new English-language learners, revealing little about what these students know and can do.

In addition, student performance is influenced by home supports, attendance, and school supports (such as class size and the availability of materials and specialist help), and it reflects the work of prior and other current teachers as well as parents and tutors as much as any individual teacher. Jackson and Bruegmann (2009) affirmed what many of us have long believed to be true –

good teachers in a school affect the performance of other teachers and, in turn, affect student achievement.

Another big concern is that standardized tests only measure a limited domain of knowledge, and miss a great many things we value. For example, science is only tested once in elementary school, and history is not tested at all. This has led to a systematic de-emphasis of these subjects in many schools, particularly those with low AYP reports. Our concerns are magnified by the fact that the standardized tests that are being used are multiple-choice tests in California. This means that many of the dimensions of learning that we should be assessing because of their importance for 21st century skill sets are not assessed. These dimensions include students' ability explain and defend their ideas in writing, to analyze carefully the research and documents they encounter, and to produce actual products based on their learning. When we have been asked as National Board candidates to cite, display, and analyze our evidence of student learning, we have discovered that standardized test scores (or any kinds of multiple choice tests) give us little or no useful information from which to draw conclusions about the effectiveness of our teaching, to make useful changes in our practice, or to convey to students and parents direction that can be used to guide new learning paths.

We do not mean to dismiss standardized test scores entirely, as we recognize they do provide some useful information. It is the nature of the information they provide that we would ask those who would use them in teacher evaluation to consider. Test scores reveal patterns of achievement at school sites and with cohorts of students in some subject areas. If the roles and contributions of teachers at a school site with cohorts of students are identified using careful analysis of scores and used to locate patterns of performance over time, they might be used as indicators to guide some direction for improvement. However, if this is to be done well, test scores must still be used together with other indicators of student achievement. We believe that test scores can be starting places for making sense of student learning if they involve careful examination of test results by teachers at a school site working with people trained to analyze test data. Teachers should use the interpretations of testing data in concert with other assessments to decide on ways to improve student performance and the teaching strategies that promote it. Done this way, the results of this careful analysis might be used as part of the evaluation of teacher performance. The questions that an evaluator or team of evaluators could use in this case might include:

- How well aligned to the instructional goals for the grade or course were the assessments the teacher used to obtain a more complete and accurate picture of student performances?
- What instructional strategies did the teacher select that matched the identified student needs and how appropriate were they in light of the goals for student learning?
- Did the reflections from the teacher seem to account for student outcomes that were directly connected to her own instruction?
- Has the teacher (alone or with his or her colleagues) selected strategies that are likely to improve her practice that match her own professional development needs to be able to deliver appropriate instruction to the particular students she is assigned to teach?

This is clearly more complicated than a simple direct link of test scores to an individual teacher as a measure of their teaching proficiency. We understand the desire for evaluation of teacher quality to be easy. We wish it could be, but teaching is complex work. Evaluating the quality of that work cannot be reduced to a simple link between the teacher and test scores. If it could, we would happily support it. However, we think that the approach we recommend honors the complexity of the work of teaching and the full range of skills that good teachers must bring to it.

Other approaches have been suggested as a way of using test scores. Some have suggested that if standardized tests are problematic for evaluative purposes, then might we consider using Value-Added Methods (VAM) to look at gains in student learning, rather than straight test scores. School leaders across the country are experimenting with VAM as a way to evaluate teachers and schools. Roughly speaking, VAM evaluates the academic growth students experience over the course of a school year, rather than comparing the current year's cohort with the previous year's. VAM also allows for adjustment of the measures for various student characteristics and for school factors. This seems like an improvement over systems that hold teachers accountable for students' average scores, or the percentage that are "proficient."

However, in a recent policy guide on the subject, Braun (2005) identified a number of specific flaws in using state tests and VAM to evaluate individual teachers. Among those flaws is the fact that students are not assigned to classes on a random basis. Also, small sample sizes in the data set, especially for individual teachers, make conclusions unreliable, and the effects of the entire school and prior school experience cannot be separated out.

Finally, as value-added methodology could raise the stakes further in an already flawed testing system, it could amplify the negative effects described above. Braun (2005) writes:

> VAM results should not serve as the sole or principal basis for making consequential decisions about teachers. There are many pitfalls to making causal attributions of teacher effectiveness on the basis of the kinds of data available from typical school districts. We still lack sufficient understanding of how seriously the different technical problems threaten the validity of such interpretations. (p. 4)

Finally, we think we need to offer a perspective about good assessment practice in real classrooms. Teachers' views of what constitutes real student learning are very different than the public perception that test scores are reliable measures of learning. We see the standardized tests as a snapshot of performance, subject to variability on any given day, that may be useful for evaluating the system as a whole. However, as an indicator of students' ability to use and apply their learning in real situations, a standardized test score is unlikely to help a teacher either advance student learning or improve their own instructional practice. For those purposes, a teacher must use assessment practices more closely linked to the actual instruction, that reveal changes in learning over time, and that allow students to demonstrate their understanding through real performance. Research on assessment has demonstrated that when teachers assess student learning in ways that help them understand both what content is being mastered and also what confusions are arising on a daily basis, the learning of students is elevated considerably, especially for those who traditionally encounter the greatest challenges. This kind of assessment is a hallmark for us of quality teaching and should be central to the teacher evaluation process.

Best Practices: Evaluation Through Professional Learning Communities

Cliff Lee, a National Board Certified Teacher and doctoral candidate at the University of California, Los Angeles, describes how ongoing evaluation from peers in a professional learning community setting can improve practice.

One of my strongest and most successful memories of teacher evaluations came from the work we did at our school from collegial support and evaluation groups. Three years ago, our professional development team (consisting of teachers and our assistant principal) created a volunteer, after-school curriculum and pedagogy brainstorm group called the Project Based Learning Cadre. The cadre was made up of teachers in different disciplines and in different grades. We met on a weekly basis, where one teacher would present a project, either in the planning or evaluation phase, and receive critical feedback through a highly-structured format. The presenter began with a brief introduction of the project, then, clarifying and specific questions from group members, followed by a whole group discussion with the presenter taking notes and finally, the presenter reflecting on the points brought up by her colleagues. In addition, each participant was asked to write up the curriculum for the project, following a highly structured format, to be saved for future use by interested teachers.

Every member raved about the benefits of this group in helping them frame, shape, evaluate, and probe deeper into their projects. The effectiveness of this collegial evaluation can be owed largely to its voluntary nature, with respectful compensation for our meeting times, write-up of our respective projects and honest, yet tactful, feedback that we received from one another.

I believe the initial buy-in with volunteers greatly enhanced the positive reaction by our cadre members, as well as by subsequent teachers that heard about it and wanted to join. In fact, some teachers were so interested in each other's projects, they took time to observe their classrooms and gave further feedback about what they noticed. I felt that we strongly held each other accountable for growth and educational outcomes because of the voluntary setup, and the positive and forward-looking framing of our work. It was also an optional professional development opportunity that served a direct benefit to our own classroom curriculum and pedagogy. We brought what we felt was most pressing and valuable to us.

The format of the presentations also allowed for greater support and constructive criticism, since there was structured space for the group to talk about the project without the presenter feeling the need to explain or even defend certain aspects of it. Although seemingly silly at first, since the group had to pretend the presenter was not in the room, this structure took the pressure off the presenter and forced the group members to ask probing and inquiring questions that led to deeper conversations.

Finally, the fact that every member would present created an egalitarian space that fostered mutual respect and trust in sharing our work. Perhaps, it may be idealistic, but I believe this communal group work of sharing curriculum and pedagogy for feedback, can be easily replicated if framed in the right manner, with an emphasis on growth and constructive criticism, rather than on punitive or judgmental evaluations. A fair compensation of the work also shows respect to the teachers and develops a level of professionalism in evaluating each others' work.

Principle 5

Evaluations should be frequent and conducted by expert evaluators, including teachers who have demonstrated expertise in working with their peers. Evaluators at each juncture should be trained in the recognition and development of teaching quality, understand how to teach in the content area of the evaluated teacher, and know the specific evaluation tools and procedures they are expected to use. Evaluations should be accompanied by useful feedback, connected to professional development opportunities, and reviewed in evaluation teams to ensure fairness, consistency and reliability.

Even if we implement our first principle and create the sophisticated tools based on what we know about good teaching, we will fail to accomplish our goal if we do not make sure that skillful evaluators are using these tools in ways that are known to be successful. Even the best tools are ineffective in the hands of those who lack knowledge and skills to use them well.

During an interview with an ACT member, Scott Laurence, a former principal in Palo Alto Unified School District, described the problems associated with evaluating teachers in content areas with which he was unfamiliar. He pointed out further disadvantages associated with his increasing distance from classroom practice. As we shared our experiences with evaluation we repeatedly surfaced the same concerns. All too often we found ourselves being evaluated by principals who had no experience teaching our grade or content area. One of us was evaluated by a principal who had only taught physical education, another was evaluated in a first grade setting by a high school physics teacher who relied on his reading about primary reading instruction to understand what the teacher was doing, still another had a principal who had been hired from the business sector on the premise that this person could manage a site more efficiently than someone with a background in education. While none of us held any animosity toward these evaluators, neither did we feel much inclined to discuss anything more than surface features of our work. Evaluation in these cases satisfied only the need to complete required paperwork.

While we are not convinced that any of the systems of teacher evaluation we examined would be the perfect fit for California, we were impressed by the level of knowledge and expertise shown by evaluators in of many of them. For example, many teachers participating in the Teacher Advancement Programs (a teacher evaluation system), were evaluated several times a year, were involved in setting their own goals for improvement, and were evaluated by people who were all trained using the same evaluation strategies. We found some of the protocols used in this program less than compelling, although we were favorably impressed by the program's use of portfolios to demonstrate professional growth and the careful way the portfolios were assessed. In Minnesota, the term "evaluation" was abandoned altogether in favor of "assessments," in which teachers' work around a particular dimension of professional practice is scored using well-defined indicators of success by teams of experts that include administrators and teachers. Danielson's *Framework for Teaching* recommends using portfolios that include both teachers' documentations of increasing skill as well as records of conversations with and observations from a teaching supervisor. Teachers are highly involved in working with supervisors in creating a professional development plan with the goal of improving practice. In every case we researched, the evaluators were highly trained in whatever approach they used and there was a high level of predictability and consistency in their approaches. Frequent observations and follow-up conversations were also common features.

There are many in the state who have already endorsed the idea that effective evaluation must include teachers. Despite our chagrin that no teachers contributed to *Students First Renewing Hope for California's Future* (Governor's Committee on Educational Excellence, 2007), the principles that we have discussed are consistent with its directives, and our recommendations will, hopefully, be carefully considered when it is time to revise and implement them. To strengthen teaching and

learning, the Committee on Educational Excellence recommended policies that will help "make teaching and education leadership true professions." Included is the notion of, "giving teachers advanced career opportunities without leaving the classroom, including mentoring and site leadership roles." With respect to opportunities to promote quality teaching, the report recommends that schools:

- Give teachers advanced career opportunities without leaving the classroom, including mentoring and site leadership roles.
- Have peers and leaders use professional
- standards and performance outcomes to evaluate teachers and principals. Let good teaching and leadership drive out bad.
- Target professional development to school priorities and student needs.
- Grant professional compensation based in part on student-performance gains, skills, and responsibilities.

(Governor's Committee on Educational Excellence, 2007, p.3)

In our discussions over the last several months, one fact that stands out is that teachers are undoubtedly less tolerant of poor performers in our profession than most of those who are empowered to evaluate us. Our conclusion is that, indeed, this would probably result in "good teaching driving out bad." But our greatest hope is that the participation of teachers would result in the construction of true professional learning communities where we hold one another accountable for helping every student succeed by learning from our individual and collective successes as well as our failures; and where schools are places in which there is no fear of making our teaching public.

Including teachers in the development of teacher evaluation systems addresses a number of important goals. Opportunities to grow and improve in the profession are precisely what's missing for too many teachers; and, as we noted earlier, teacher retention is negatively affected by those missing opportunities. Ambitious teachers who have thrived in the classroom are often drawn out of teaching if they want to pursue greater challenges or the means to exert greater influence on the direction of public education. LAUSD National Board Certified Teacher Jane Fung adds the voice of personal experience: "I have seen good, newer teachers leave the classroom each year in search of another place or career that will recognize and value their efforts more than classroom teaching does."

Here we see additional benefits to teacher involvement in designing and implementing a better evaluation program. An evaluation system with teachers at the center will make evaluations more credible, more productive, and more valuable to teachers. Schools are more likely to retain skilled teachers who have opportunities to apply their expertise in work with colleagues. Teachers who take on these expanded roles benefit from the process in equal measure with their peers, expanding their own knowledge, skills, and perspectives and becoming better able to take the additional steps to promote quality work.

An example of this kind of expansion of professional knowledge and skill comes from the National Board Resource Center at Stanford. Support providers for National Board candidates are frequently recruited for the center's support program from National Board Certified Teachers who have already been through the program. Year after year they attest to the powerful impact on their own teaching as they help others learn to align their practice to National Board standards. A typical comment is, "As hard as it is to get out of bed on Saturday mornings, it is so worth the effort. At the end of the day, I always leave a better teacher than when I came in."

While teacher evaluation should primarily be about recognizing and improving the quality of teaching, we acknowledge that we must consider the serious decisions that must be made through the evaluation process. The opportunity to advance along the continuum of professional growth is

one of the options that is most important to us as teachers. It is equally important to make sure that a fair system is in place to determine which teachers remain in the profession.

Principle 6

The evaluations that lead to teacher tenure must include more extensive evidence of quality teaching. This evidence should be collected and reviewed by both the teacher and trained evaluators and should include documentation that shows that the teacher's practice exhibits the standards that exemplify good practice. The process should be an ongoing part of a serious teaching induction process that helps novices grow in their profession, with the help of mentors and coaches, guided by clear standards of practice.

> Teachers indicate that the most obvious technique used to assess teacher quality – the formal observation and evaluation is not doing the job. In fact, only 26% of teachers report that their own most recent formal evaluation was "useful and effective." The plurality – 41% – say it was "just a formality," while another 32% say at best it was "well-intentioned but not particularly helpful" to their teaching practice. Almost 7 in 10 teachers (69%) say that when they hear a teacher at their school has been awarded tenure, they think that it's "just a formality – it has very little to do with whether a teacher is good or not."
>
> – Duffet et al. (2008, p. 3)

Awarding tenure to teachers who merit it could be done with greater certainty if we created an evaluation system that was a more reliable indicator of quality. Using elements that are already in place, teachers in collaboration with evaluators could compile bodies of evidence that document how their teaching meets standards for practice. Adding to the portfolio begun in pre-service, teachers would be able to show how they have continued to align their practice to ever more sophisticated standards based on the continuum of professional practice we have recommended. A logical and important way to do this would be to build on PACT and TPA and design a similar assessment for moving from a preliminary to a clear credential that reflects the growth in skills and knowledge a teacher has mastered during the induction period.

This assumes, of course, that PACT and TPA are developed to meet their full potential. This means that these assessments are uniformly implemented, scored by panels of knowledgeable assessors, and used to recommend an induction program that is focused to ensure that novice teachers have every opportunity and support available to learn and get feedback about their work. We would further recommend that BTSA be improved to incorporate some of the same features we believe are integral to a good evaluation system – mentoring and coaching by accomplished peers who have experience in the same content areas as the novice teachers with whom they work and, where possible, who teach in the same school.

Awarding permanent status along with dismissal of chronically poor teachers are issues on which we spent a significant amount of time. We have a large stake in making sure that our colleagues not only contribute effectively to students' learning but also contribute to our collective professional knowledge. We all want to support and contribute to the learning and growth of our novice peers, but we want to do so as part of a system that truly believes that communities of teachers should hold one another accountable. Very few teachers are denied tenure after they have completed their induction programs even though we, who have mentored them or worked with them as grade level or department colleagues, may have concerns about granting them permanent status. The fact is, though, the observations and work done in BTSA or in collaborative work are not usually considered in tenure evaluations. Tenure approval is usually based solely on the recommendation of the principal.

Because we all have a large stake in making sure that our colleagues are ready to work effectively with students and be a productive part of school learning communities, it seems fitting that we should have a role in determining permanent status. Tammie Adams, a union representative and National Board Certified Teacher from the Oakland Unified School District, posed the question to another union representative about how a community of teachers should exert influence on tenure decisions. He voiced the same concerns about the cavalier manner in which tenure decisions are often made. While he did not come down on the side of a more rigorous summative evaluation before tenure could be granted, he did voice a strong opinion about roles teachers should play in determining who is awarded tenure, and the degree of tolerance we should have for novices (and others) about whom we have serious concerns. A benchmark summative assessment at the end of the probationary period, created with significant involvement of accomplished teachers and supported by unions and professional communities, could go far in ensuring that good decisions are made with respect to who stays in the profession.

Principle 7

Evaluations should be accompanied by useful feedback, connected to professional development opportunities, and reviewed in evaluation teams to ensure fairness, consistency and reliability.

> I'd like to see the teacher evaluation process become meaningful in terms of teacher growth. I've never seen on any of my evaluations a suggested area that I might explore more deeply in my instruction. Why is that? Is it that the evaluators have nothing to suggest, don't know what to suggest, or don't bother to take the time to actually analyze my instruction?
> – Kathie Marshall, Los Angeles Unified School District

Good teaching practice involves offering frequent feedback to students about their progress toward well-understood performance goals. The same principle should apply to teacher evaluation to improve the quality of practice. Teachers need to be helped to internalize the vision of good teaching as exemplified by standards and they must be helped to hold their own practice up for comparison to those standards. Conversations that help teachers to reflect on their instructional decisions, the rationale for making them, the results of their actions, and their impact on students are essential if we are to help teachers advance the quality of their practice throughout their professional lives. The kind of feedback we are talking about is not a simple checklist of the sort commonly employed in the current evaluation protocol nor is it a series of comments on the implementation of a scripted curriculum. Rather, it is a conversation conducted by highly skilled and knowledgeable evaluators who understand how to help candidates reflect deeply and who know how to ask questions that push teachers to see their own practice more clearly and to engage with evaluators (and, by extension, other colleagues) in addressing the challenges of their own classrooms.

We were enthusiastic about the work done by teachers in Minneapolis where peer evaluators helped teachers identify growth areas based on evaluations and then helped them select coursework specifically designed by and for teachers in the district that aligned with identified areas of work. In this program, teachers who attended classes taught by other teachers with expertise in the targeted skill areas were found to incorporate their learning in subsequent evaluations and to create projects that demonstrated mastery of the target performance area. These projects were not just done to fulfill a course requirement, as is the case in so many continuing professional development courses taken by teachers, they were used to apply or analyze classroom practice and were evaluated and scored by an evaluation team according to a rubric based on the standards. Teachers often submitted artifacts of their teaching practice that showed changes they had made as a result of their

coursework only to be given additional feedback to support continued work toward mastery in exactly the same way that teachers ask students to revise and refine their work.

An important component of providing feedback is ensuring that evaluators have the support and skill to provide it. This means that they need to be able to talk with other evaluators about the dilemmas they encounter in providing support and they must have opportunities to frequently "re-calibrate" assessments of what constitutes the target level of performance for each stage of teacher development. Those of us who have worked on holistic scoring of various kinds of assessments, including those of our National Board candidates, know that this process increases our ability to engage in good collaborative reflection that improves our own practice as well as that of those we assess. This aspect of the evaluation system we recommend becomes an important vehicle for changing the culture for teaching in schools by promoting shared accountability and changing the nature of professional conversations.

References

Barnes, G., Crowe, E., & Schaefer, B. (2007). *The cost of teacher turnover in five school districts: A pilot study.* Washington, DC: National Commission on Teaching and America's Future. http://www.nctaf.org/resources/demonstration_projects/turnover/documents/CTTFul lReportfinal.pdf.

Braun, H. (2005). *Using student progress to evaluate teachers: A primer on value-added models.* Princeton, NJ: Educational Testing Service.

Danielson, C. (1996). *Enhancing professional practice: A framework for teaching.* Alexandria, VA: Association for Supervision and Curriculum Development.

Darling-Hammond, L. (2007). *Recognizing and enhancing teacher effectiveness: A policymaker's guide.* Washington, DC: Council of Chief State Schools Officers.

Darling-Hammond, L., Wilhoit, G., & Pittenger, L. (2014). Accountability for college and career readiness: Developing a new paradigm. Education Policy Analysis Archives, 22(86). http://dx.doi.org/10.14507/epaa.v22n86.2014.

Duffett, A., Farkas, S., Rothertham, A. J., Silva, E. (2008). *Waiting to be won over: Teachers speak on the profession, unions, and reform.* Washington, DC: Education Sector.

Governor's Committee on Educational Excellence. Summary. (2007). *Students first: Renewing hope for California's future.* Sacramento, CA: Author.

Jackson, C. K., & Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *American Economic Journal: Applied Economics 1*(4*)*, 85-108.

Shulman, L. (1988). A union of insufficiencies: Strategies for teacher assessment in a period of educational reform. *Educational Leadership, 46*, 36-41.

Sterling, D. R. (2004). The teacher shortage: national trends for science and mathematics. *The Journal of Mathematics and Science: Collaborative Explorations, 7*, 85-96.

Toch, T., & Rothman, R. (2008). *Rush to judgment: Teacher evaluation in public education.* Washington, DC: Education Sector.

About the Authors

Accomplished California Teachers*
scope@stanford.edu
*The original report, *A Quality Teacher in Every Classroom: Creating a Teacher Evaluation System that Works for California,* was written in 2010 by a team of California-based teacher leaders, who were part of the Accomplished California Teacher (ACT) network operating under the auspices of the National Board Resource Center at Stanford University.

About the Guest Series Editor

Dr. Linda Darling-Hammond
Guest Series Editor
Stanford University
ldh@stanford.edu
Linda Darling-Hammond is Charles E. Ducommun Professor of Education at Stanford University where she is Faculty Director of the Stanford Center for Opportunity Policy in Education. Her latest book is *Beyond the Bubble Test: How Performance Assessments Support 21st Century Learning* (Wiley, 2014).

Please contribute commentaries at http://epaa.info/wordpress/ and send errata notes to Gustavo E. Fischman fischman@asu.edu

Join EPAA's Facebook community at https://www.facebook.com/EPAAAAPE and **Twitter feed** @epaa_aape.

arquivos analíticos de políticas educativas
conselho editorial