

BEYOND SATISFACTORY: REDEFINING TEACHER EVALUATION AND SUPPORT TO IMPROVE TEACHING AND LEARNING



The Education Trust—West

The Education Trust–West works for the high academic achievement of all students at all levels, pre-K through college. We expose opportunity and achievement gaps that separate students of color and low-income students from other youth, and we identify and advocate for the strategies that will forever close those gaps.



The Education Trust–West

1814 Franklin St., Suite 220, Oakland, Calif. 94612

T 510/465-6444 • F 510/465-0589

www.edtrustwest.org

Executive Summary

Improving teacher quality is one of the most significant investments we can make to ensure educational achievement for all students, but especially for students of color and those living in poverty. While teacher evaluation policies and practices ought to help educators assess teacher effectiveness and inform the supports offered to help teachers improve, California's teacher evaluation laws are ineffective and limit educators' efforts to do either. Several pioneering school systems—including both school districts and charter management organizations—are trying to change this.

In this report, we present our findings from the first year of a two-year study examining the perspectives of teachers and school leaders in seven school systems piloting new ways to evaluate and support teachers. We studied both districts and charter management organizations (CMOs) that were either piloting or newly implementing a multiple-measure teacher evaluation system in some or all of their schools in the 2012–13 school year. The following is a summary of our key findings.

KEY FINDINGS

1. Both administrators and teachers agree the effort to replace their old teacher evaluation process with a new one based on multiple measures of teaching quality was a lot of work but worth it.
2. Teachers and administrators express enthusiasm about having clear and detailed expectations for effective teaching that anchor the teacher evaluation process and foster a shared language and an understanding of teaching and learning.
3. There is strong backing for the use of multiple measures of teacher practice, but support varies for each individual measure—with the most confusion about the student achievement growth measures, especially in non-tested grades and content areas.
4. Teachers believe they are being observed more regularly, that the feedback they are receiving is more meaningful, and that the evaluation process has created more opportunities for them to think about and to improve their classroom practices.
5. Despite having some initial misgivings, most teachers are experiencing evaluation as a vehicle for improvement, not as a “gotcha;” and their supports are becoming both more plentiful and better aligned with the evaluation criteria over time.
6. Training for both teachers and evaluators¹ is absolutely critical. By working hard to ensure broad understanding of the evaluation framework and, in some cases, requiring evaluators to “certify” their competence, these school systems have created broadly shared understandings of and expectations for specific characteristics of good teaching.
7. Virtually every school system made significant changes in its teacher evaluation and support process along the way; indeed, the willingness to do so as issues arose helped build trust and confidence.

State-level policymakers and district leaders need to acknowledge that both our state regulatory framework and most district teacher evaluation systems undermine efforts to improve teaching and learning in California schools. When evaluation processes function at all, they basically tell all our teachers that they are just fine—or “satisfactory.” But this

does not make sense when fewer than one out of three of our fourth and eighth grade students is proficient in reading and mathematics, and nearly half of our college freshmen are not prepared to attend college without remediation.

Fortunately, some of our own school districts and CMOs have been leading the way to fix this problem. Collectively, these school systems are redefining the teaching profession. They have moved from an old model, in which teachers taught mainly in isolation and earned salary increases for each year they taught regardless of their performance, toward a new system focused on student growth and professional collaboration. We hope that their experiences will urge state and local leaders to act on the following policy and practice recommendations.

STATE-LEVEL RECOMMENDATIONS:

1. Strengthen state guidelines and requirements for teacher evaluation.
2. Remove barriers to staffing decisions that are in the best interests of students, with particular emphasis on high-need students.
3. Leverage funding sources to support local teacher evaluation work.
4. Require districts to report staffing data.
5. Develop programs and policies to place and retain the best teachers in our highest-need schools.

LOCAL-LEVEL RECOMMENDATIONS:

6. Make teacher evaluation and support work a top priority.
7. Ensure all key stakeholders are engaged.
8. Adopt or create a framework of quality teaching with clear indicators that are aligned to standards for student learning.
9. Establish multiple measures of teaching quality and make sure they are valid and reliable.
10. Train people thoroughly on both the evaluation indicators and process.
11. Establish a robust certification process for evaluators and ensure they are empowered to meaningfully differentiate teacher effectiveness.
12. Develop a detailed plan for all stages of implementation, including a pilot phase.
13. Identify structures and resources for instructional support in advance.
14. Pay particular attention to principals.
15. Accept that the system will take time to fine-tune.

Teachers deserve honest feedback and support to become ever more effective. And now is the perfect time: Now that we've adopted a new set of standards for our K–12 students—the Common Core State Standards in English language arts and mathematics, Next Generation Science Standards, and English Language Development standards—we need to bring our standards for teaching into line. Our students deserve nothing less.

Beyond Satisfactory: Redefining Teacher Evaluation and Support to Improve Teaching and Learning

BY JEANNETTE LAFORS

Improving the quality of teaching is one of the most significant investments we can make to ensure educational achievement for all of our students. Nobody who cares about the quality of education in California can afford to ignore it.

This issue is especially salient for us at The Education Trust–West because quality teaching is particularly critical for students of color and those living in poverty, who typically lack opportunities afforded their more privileged peers.² While all students deserve to have effective teachers, struggling students especially need teachers who can ensure they grow academically and achieve at high levels. And yet high-need students—most often low-income students, English learners, and students of color—are disproportionately assigned less effective teachers, perpetuating patterns of low achievement among our Latino, African American, English-learner, and low-income student populations.³

In recent decades, our state has taken steps to improve the quality of teaching. We’ve designed and put into place a support system for beginning teachers that is considered a national model. We’ve developed a framework for teaching quality,⁴ established a peer assistance and review program to support veteran teachers with unsatisfactory ratings, and required K–12 teacher candidates to pass a teaching performance assessment in order to earn their preliminary teaching credential.

But we have so far been unwilling to tackle a problem that diminishes the value of these other improvement efforts: we have an evaluation system that, when it functions at all, basically tells our teachers that they are all just fine. All teachers—great, good, and poor—are denied feedback they both need and deserve as professionals to grow and improve. If we are serious about improving teaching, this needs to change.

Instead of undercutting the effectiveness of other efforts, teacher evaluation needs to become an integral part of a system of

quality assurances for effective teaching. That system starts with establishing strong selection criteria for aspiring teachers and continues with providing high quality teacher training, ensuring teacher candidates are licensed only when they have met rigorous criteria, providing growth opportunities for all teachers, and incentivizing teachers who master their craft to stay in the profession—and to teach where they are most needed.

All professionals need honest feedback and tailored help to get better. Teachers are no different. Our teacher evaluation policies and practices ought to assure that all teachers are regularly and honestly assessed, and that they get help in improving any weaknesses identified in that process.

Most of our teachers, however, get neither the regular feedback nor the supports they crave and deserve. The current law governing teacher evaluation in California—the Stull Act—mandates annual evaluations only for non-tenured teachers. And, though improving student learning is the most important role teachers play, the Act does not require that evidence of student growth be incorporated into a teacher’s evaluation rating in a meaningful way, or that feedback from parents or students be considered.

If our students were all doing just fine, this might be tolerable. But they are not doing just fine. In both the fourth and eighth grades, fewer than one out of three students in California is proficient in reading and mathematics,⁵ and nearly half of California’s college freshmen are not prepared to attend college without remediation.⁶ With results this low—far lower for every group of children than in other states—how can nearly all of our state’s teachers be rated as “satisfactory”?

Jeannette LaFors, Ph.D., is Director of Equity Initiatives at The Education Trust–West.

We have got to get serious about fixing this problem. Other states have. In the last several years alone, 35 states have moved to require that student achievement growth be a significant factor in teacher evaluations, and 28 states require annual evaluations of all teachers.⁷

Fortunately, we don't have to look outside of California for some good models, for some of our own school districts and charter school management organizations have been leading the way. We at The Education Trust–West have had the good fortune to be able to study the early stages of work in seven pioneering school organizations in California that are implementing multiple-measure teacher evaluation and support systems designed to improve teacher effectiveness and raise student achievement.

Collectively, these systems are redefining the teaching profession. They have moved from an old model, in which teachers taught mainly in isolation and earned salary increases for each year they taught regardless of their performance, toward a new system focused on student growth and professional collaboration. We hope that their experiences will inspire state and local school leaders to take on this important work.

About Our Research

This study was designed to capture how teachers and school leaders adopting innovative, multi-measure teacher evaluation and support systems have experienced and responded to initial efforts to build coherent evaluation and support structures. In addition, we sought to surface promising or best practices emerging from the new teacher evaluation models that could inform other California districts, charter management organizations, and state policymakers intent on improving teacher evaluation.

In 2012–13 we visited teachers and leaders in 22 schools across seven school systems⁸ throughout the state: three K–12 school districts of varying size, region, and demographic make-up, and four CMOs serving high-need students, primarily in Los Angeles, but also the Bay Area, and Central Valley. The three districts are: Los Angeles Unified, Lucia Mar Unified, and Oakland Unified. The four CMOs are: Alliance College-Ready Public Schools, Aspire Public Schools, Green Dot Public Schools, and Partnerships to Uplift Communities (PUC) Schools—which together compose a group known as The College Ready Promise (TCRP). Table 1 indicates several characteristics of each school system.

We gathered and reviewed more than 300 artifacts, interviewed roughly 200 teachers and administrators, observed over 100 classrooms, and surveyed over 1,000 educators. For more details, see the data collection sidebar on p. 4.

TABLE 1: Key characteristics of school systems

System	Location	Size (2012–13 data)	% Free- and Reduced-price Meals	% Students of Color	# of Teachers	Unionized Teachers (Yes/No)	Key Feature
LOS ANGELES UNIFIED	Los Angeles County	>660,000 students in 1,100 schools (K–12)	80%	90%	33,000	Yes	Initial pilot of 70+ schools led to major streamlining effort
LUCIA MAR UNIFIED	San Luis Obispo County	>11,000 students in 17 schools (7 TAP schools)* (K–8)	49%	47%	500	Yes	Master teachers serve on school leadership team and deliver weekly PD & coaching
OAKLAND UNIFIED	Alameda County	>47,000 students in 86 schools (K–12)	81%	91%	2,700	Yes	SIG** schools piloted new teacher evaluation process with help from Aspire
ALLIANCE	Los Angeles County	10,000 students in 21 schools (6–12)	93%	100%	430	No	Built leadership capacity by certifying teacher leaders as evaluators
ASPIRE	Bay Area, Central Valley, Los Angeles***	12,600 students in 34 schools (K–12)	80%	>85%	628	No	Data-driven system has powerful teacher supports
GREEN DOT	Los Angeles County	10,400 students in 18 schools (6–12)	96%	99%	519	Yes	Expanded coaching supports are based on initial feedback
PUC	Los Angeles County	4,200 students in 13 schools (K–12)	100%	99%	213	No	Culture of revision accommodates bottom-up changes

Source: California Department of Education, 2012–13 data and interview data.

*TAP was piloted initially in seven schools. “TAP” is shorthand for “The System for Teacher and Student Advancement” managed by the National Institute for Excellence in Teaching.

**School Improvement Grant

***Includes schools in the following counties: Alameda, Los Angeles, Sacramento, San Joaquin, San Mateo and Stanislaus.

DATA COLLECTION

Focus Groups and Interviews

We conducted 21 teacher focus groups including more than 80 teachers across different grade levels and content areas, and we ran 13 evaluator/observer focus groups with more than 50 individuals. We also conducted four student focus groups at different secondary schools and one parent focus group. In addition, we performed nearly 40 semi-structured interviews with principals and system leaders.

Observations

Across 16 school sites, we visited over 100 classrooms and observed more than a dozen meetings (e.g., professional development sessions, post-observation debrief meetings, and leadership team meetings).

Artifacts

We reviewed more than 300 documents across all seven systems, including teacher evaluation handbooks, system data analyses, newsletters, professional development materials, teacher lessons, and online resources.

Educator Survey

During the 2013 spring semester, we surveyed 1,035 educators across six of the districts and CMOs—905 teachers and 130 individuals who have responsibilities related to teacher support, development, mentoring, coaching, observation, and/or evaluation (e.g., administrators, teacher leaders, and teacher coaches). The overall response rate was greater than 50 percent, ranging by system from 37 to 56 percent.⁹ While we did not survey educators in Los Angeles Unified with our instrument, we accessed their survey reports and highlight key findings throughout this report.

Our Findings

While each of the seven school systems approached the redesign of evaluation in a different way, including how they leveraged resources to support that work,¹⁰ we found many commonalities across those systems in how educators experienced the changes. This report summarizes what we learned in a set of seven findings, illustrated with spotlights to provide detail on what districts and CMOs are doing to improve student outcomes.

Finding 1: Both teachers and administrators agree that the effort to replace their old teacher evaluation process with a new one based on multiple measures of teaching quality was a lot of work but worth it.

Implementing a new teacher evaluation system was challenging work. It took time and effort for participants to gain a deep understanding of the criteria for effective teaching, and it was hard to master all the other pieces of the system so quickly. Indeed, many educators likened the rollout of their teacher evaluation system to “building the airplane as we’re flying it.” They grappled with putting so many new things in place and struggled with several of them, including frustration with cumbersome online tools, understanding the role of student achievement measures in teachers’ evaluation ratings, and uncertainty about how the ratings might affect things like salary or tenure.

Nevertheless, when we asked teachers and administrators whether, in hindsight, they would choose to adopt their new teacher evaluation system or not, nearly every one of them stated they would.

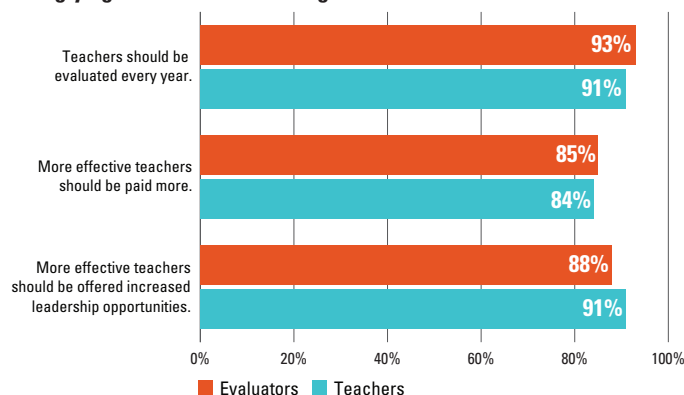
Teachers for the most part asserted that previous evaluation experiences lacked depth and relevance, and never did much to help them improve their teaching. In contrast, their new evaluation and support system was helping to transform teaching and learning in their classrooms and throughout the school. As one teacher shared, “I am so excited about what a better teacher I am now. ... My job is so much easier. I have better practices, better management skills. I have someone I can go to when I teach a lesson, to ask [how I can] make it more understandable, and they share ideas with me. Everybody is working for everyone else to try and make it the best teaching experience for us and the kids ... for every single lesson.”

Like teachers, administrators acknowledged that the old process simply did not work and that the new process was offering more to teachers. “How we were evaluating teachers in the past was so superficial “and subjective,” one administrator said. “Teachers were ‘competent’ or ‘incompetent’—there was no in between. This process allows teachers to hear feedback about their teaching and receive quality suggestions for how to improve.” Even though improving teacher evaluation and professional supports places intensive demands on them, principals remain committed to the process because of the tangible improvements in classroom instruction they see.

Survey results indicate strong support for the process: Three-quarters of teachers noted the evaluation process was useful to them,¹¹ 92 percent of teachers and 94 percent of evaluators agreed that teachers should be evaluated every year. Moreover, most teachers viewed the evaluation system as a way to reward teaching effectiveness and to open up career opportunities. Specifically, more than 80 percent of teachers

and evaluators surveyed believe that more effective teachers should be paid more and roughly 90 percent of teachers and evaluators believe more effective teachers ought to be offered increased leadership opportunities.¹² (See Figure 1).

FIGURE 1: Percent of evaluators and teachers who agree and strongly agree with the following statements



Source: ETW survey data, spring 2013.

Finding 2: Teachers and administrators express enthusiasm about having clear and detailed expectations for effective teaching that anchor the teacher evaluation process and foster a shared language and an understanding of teaching and learning.

Each of the evaluation systems we studied is anchored by a description of effective teaching that includes four or five domains, each accompanied by a brief profile of what effective and ineffective practice looks like:¹³

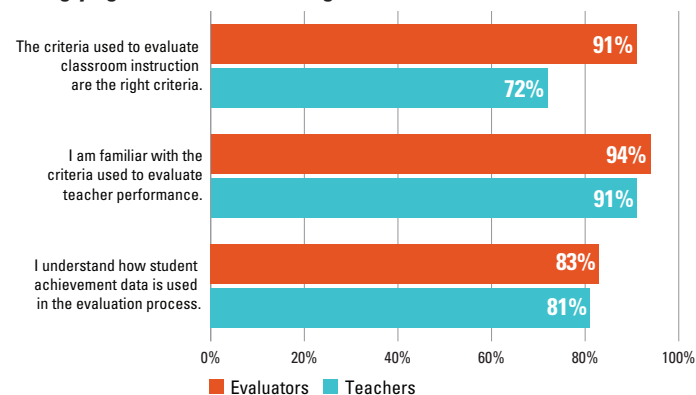
- lesson planning
- assessing student learning
- creating a positive classroom learning environment
- using research-based teaching methods
- demonstrating professional responsibilities and growth

Each criterion—informed by Charlotte Danielson’s research-based Framework for Teaching¹⁴ and other frameworks—is presented on an observation rubric. There, educators note distinctions between different levels of performance. A Level One rating would be considered “emerging” or “unsatisfactory” while a Level Four or Level Five rating might be considered “mastery” or “exemplary.”

Most educators interviewed in this study agreed that the criteria used in their system were clear and contained detailed descriptions. In addition, both evaluators and teachers believed that the criteria used to evaluate teacher performance were the right criteria, though evaluators communicated greater agreement than teachers did (91 and 72 percent, respectively). Almost all educators said they were familiar with the criteria used to evaluate teacher performance, though fewer

educators fully understood how student achievement data were used in the evaluation process (See Figure 2). Separate surveys conducted by Los Angeles Unified yielded similar findings.^{15, 16}

FIGURE 2: Percent of evaluators and teachers who agree and strongly agree with the following statements



Source: ETW survey data, spring 2013.

Teachers, teacher coaches, and administrators alike noted how the rubrics facilitated the creation of a shared language among educators about effective teaching. They also indicated that the frameworks and rubrics supported more evidence-based conversations about teaching and learning. Several teachers noted that the conversations they had with their evaluator were more objective than those they had experienced in the past. This objectivity is achieved in part because evaluators capture in writing what both the teacher and students are doing in class during a classroom observation and then tag those notes as evidence aligned to specific elements of the instructional rubric. One teacher said, “We went over my four goals, and it was very objective. ... It was a conversation about evidence. ... If I scored a 3 [on the rubric] we looked at what a 4 would require. It felt very natural and objective and low-stress.”

Finding 3: There is strong backing for the use of multiple measures of teacher practice, but support varies for each individual measure—with the most confusion about the student achievement growth measures, especially in non-tested grades and content areas.

“It helps us look at our instruction from different points of view, including what our students need.”
—Teacher



SYSTEM SPOTLIGHT: THE COLLEGE READY PROMISE

The College Ready Promise (TCRP) is a nonprofit consortium of four CMOs¹⁷ serving nearly 40,000 predominantly low-income students that came together in 2009 under a grant funded by the Bill & Melinda Gates Foundation to achieve a shared goal: that all their teachers be “supported to become highly effective in order to ensure all students achieve at high levels and graduate college-ready.” They outlined five things they needed to accomplish this goal, the first of which was to establish a common language and definition for highly effective teaching in a college-ready culture embodied in a framework that sets clear performance expectations. They partnered with Teaching and Learning Solutions to design a unique teacher observation rubric based on nationally accepted observation standards. Originally identifying indicators and standards of effective teaching across six domains, TCRP member organizations narrowed their framework over numerous feedback sessions and revisions to four domains: (1) data-driven planning and assessment, (2) classroom learning environment, (3) instruction, and (4) professional responsibilities. As one principal said, “[The rubric] gives me a solid foundation on what good instruction is. It gives me a focus for my conversation with teachers. I’ve become a better instructional leader through this process.” Another administrator described how much educators were embracing the rubric: “Our teachers and principals absolutely love it. They use it, they talk about it. It’s driving practice, reflection, and coaching, and it’s transparently inherent in the conversations about instruction.” Close to 90 percent of TCRP teachers and evaluators agreed they were familiar with the criteria used to evaluate teacher performance.

As the rubric was being developed, TCRP approved the use of student growth percentiles—a measure comparing the growth a teacher’s students make to that of other students who start at the same achievement level—to evaluate student growth as one of the indicators of a teacher’s overall measure of effectiveness. The group also established processes for creating robust teacher and student data systems so teachers could have real-time access to their data. Teachers and administrators were included in the process of determining weights for each component of the teacher evaluation; individuals interviewed across the CMOs spoke highly of the opportunities to provide feedback on the evaluation system design.

In each of the teacher evaluation and support systems we studied, teachers’ evaluation ratings are dependent on data from multiple sources, which also inform professional development offered to teachers. In general, there are three main sources of data in each of the systems (see Appendix A for details that include variations on the weights of these measures):

- Observations of classroom practice (formal and informal)
- Student growth data (individual and/or school-wide student growth)
- Stakeholder feedback (from family, student, and/or peer surveys)

Most individuals we spoke with found their multi-measure teacher evaluation system to be more balanced and objective than any previous system. Teachers generally appreciated having multiple perspectives inform their evaluation, and cited them as levers for improvement. “It helps us look at our instruction from different points of view, including what our students need. It’s definitely helped me improve instruction,” one teacher said.

Furthermore, various data sources related to instructional practices guide professional development decisions and resource allocations across systems. According to both teachers and administrators, survey data from parents, students, and teachers, student achievement results, and observation data are all used to target professional development supports not only for individual teachers, but also for teams of teachers or an entire school.

While educators generally value the use of multiple sources of evidence, they also identified challenges and concerns with

various measures. For classroom observations specifically, many teachers preferred more frequent, shorter, unannounced observations compared with one or two longer, scheduled formal observations. They argued a formal observation might not accurately represent their day-to-day teaching—either because the formal lesson was designed to “showcase” so many instructional strategies or because something unexpected might happen during their lesson to dramatically affect their plan and therefore their rating. While many evaluators also see the benefit in frequent observations, they also valued knowing how well, given notice, teachers could craft a powerful lesson demonstrating their mastery of a full range of teaching strategies that address the learning needs of all their students. And while over 80 percent of teachers across all seven systems indicated that they trusted their evaluator to evaluate them objectively,¹⁸ we heard teachers readily note that their observation ratings depended on who evaluated them.

When it came to measures of student growth, teachers generally did not shy away from a responsibility to show evidence of student learning, but they expressed concerns about how accurately the California Standards Tests captured student learning, how future assessments would be able to do so, and how to most fairly assess teachers who have students without statewide test scores.¹⁹ Many of the districts and CMOs have begun to address these questions and have developed alternative means to assess student achievement and growth as California transitions from one set of assessments to another, but even those solutions pose challenges.

Some teachers whose evaluation includes student feedback worried about whether younger students fully understood

questions on the survey, and secondary teachers raised issues about student sampling, voicing concerns that surveying some of their students rather than all of them might skew the results. Many teachers see peer surveys as problematic, in part because they did not always feel they had adequate opportunities to observe their colleagues in all the areas they were asked to rate them, but also because they were reluctant to give their colleagues ratings that could have a negative impact on their salaries.

All in all, despite the issues teachers and evaluators encountered with a multiple-measure evaluation system, they supported the concept of gathering several indicators to determine teaching effectiveness.

WHAT KIND OF SURVEYS INFORM TEACHER RATINGS?

Survey design is deceptively complicated. Given the high stakes involved in teacher evaluation, schools and districts may want to think twice before writing their own survey questions. By selecting a well-tested survey or working with a trusted research team to design a new one, districts can ensure that survey questions will accurately and dependably measure teacher effectiveness and correspond with student results. The Tripod student survey developed by Dr. Ronald Ferguson,²⁰ for example, asks students multiple questions to get at a number of themes, and the results from this survey strongly predict student learning.

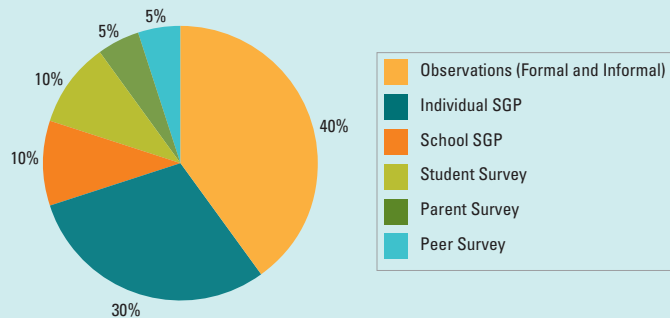


SYSTEM SPOTLIGHT: ASPIRE PUBLIC SCHOOLS

Aspire Public Schools’ teacher evaluation and support model incorporates multiple measures, including observations (both formal observations and informal mini-observations), the teacher’s individual student growth percentile data,²¹ school-wide SGP data,²² and feedback generated through student, parent, and peer surveys.

For teachers with students assessed with California Standards Tests (grades 3–11), Aspire determined, with input from educators in its system, the following weights for each of these measures, as indicated in Figure 3 below:

Figure 3: Teacher evaluation measures for Aspire Public Schools, 2012–2013



Source: Aspire Public Schools artifacts and interview data.

For teachers with students who do not take CSTs, Aspire gives more weight to the school-wide SGP while it pursues alternative means of assessing student growth in a particular grade and a specific content area.

The evidenced-based approach to teacher evaluation and support is valued at Aspire, and evaluators are trusted as credible. Specifically, 99 percent of Aspire teachers surveyed reported they were encouraged to use evidence of student learning to inform practice, and 84 percent of teachers reported that they trust their evaluator to assess them objectively. While there was a discrepancy between

how fair teachers and evaluators reported the evaluation process to be (75 versus 90 percent, respectively), 96 percent of teachers stated they understand how they are (or will be) evaluated, and 95 percent of teachers agreed (along with 97 percent of evaluators) that teachers should be evaluated every year.

As one teacher said, “It’s never a ‘gotcha,’” and other teachers articulated their experiences with the teacher evaluation process as objective. One teacher stated, “The conversations are powerful because it’s not what one person’s interpretation of what effective teaching is,” while another teacher noted, “My debrief was about evidence and reflection.”

Like other CMOs in The College Ready Promise consortium, Aspire implements a rigorous evaluator certification process for principals. If a principal does not meet the certification criteria, he or she must work with someone who is certified during the teacher evaluation process.

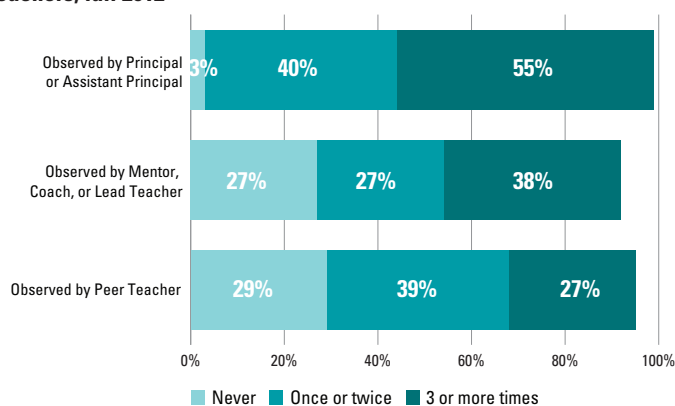
In addition to acknowledging the power of multiple measures as part of their evaluation, teachers at Aspire frequently described ways that the evaluations influenced the professional development they received. Every teacher is expected to set individual improvement goals in addition to working on team-specific and school-wide goals determined by students’ needs. Of the Aspire teachers surveyed, 80 percent agreed that the evaluation process was useful to them and that professional development at their school is useful to their teaching, with 70 percent also agreeing that professional development at their school addressed their individual needs. Teachers who have worked in other systems compared how frequently they were observed, coached, and supported in those schools with their experiences at Aspire. One teacher said, “Compared to a traditional district [where I used to work], the steps that Aspire has made to support teachers is really differentiated. ... You’re getting specific support based on your needs.” Teachers affirmed generally what one teacher put to words: “The most important thing we’re going to get from this [evaluation system] is how to improve our practice.” And finally, some teachers voiced their approval of what they considered to be a system that affirmed, even incentivized, being an effective teacher committed to classroom teaching.

Finding 4: Teachers believe they are being observed more regularly, that the feedback they are receiving is more meaningful, and that the evaluation process has created more opportunities for them to think about and to improve their classroom practices.

While the frequency and nature of observations differed by system,²³ teachers across systems noted that they receive more frequent and meaningful feedback than they have in the past, along with more opportunities to reflect upon and improve their teaching.

The survey data revealed that more than 95 percent of teachers were observed by an administrator at least once in the fall of 2012, but more than half of the teachers (55 percent) were observed more than three times. In addition, nearly 70 percent of teachers were observed by a coach, mentor, or lead teacher that same semester, and almost a third had more than three such observations. The same was true for peer observations (See Figure 4).

FIGURE 4: Frequency of observations, by observer role, reported by teachers, fall 2012

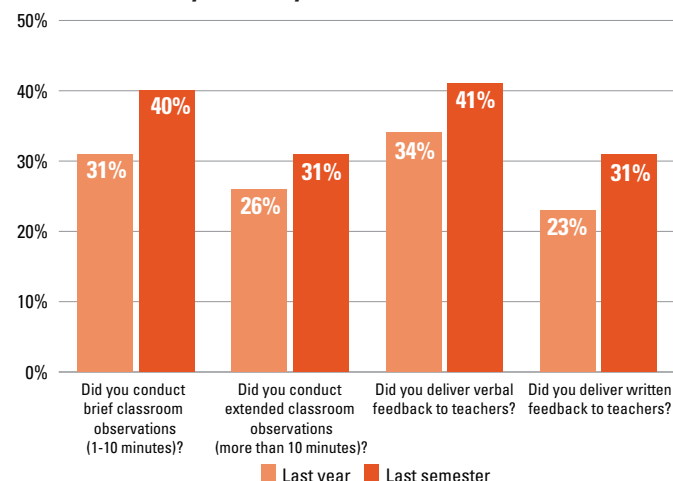


Source: ETW survey data, spring 2013.
Total percents may not add up to 100% because of n/a responses.

In response to educator feedback and current research, some systems in our study began to experiment with increasing the number of informal observations in lieu of the more traditional two-formal-evaluations-per-year models.²⁴ For example, in Green Dot's "many mini model," a pilot initiated in spring 2013, participating teachers had six unannounced observations by two administrators, with each administrator observing the teacher a minimum of two times. In the same vein, PUC Schools shifted from two formal observations to one formal spring observation enhanced by several informal observations conducted throughout the year.

Administrators indicated that they conducted both extended and brief observations more frequently in fall 2012 than they had the previous academic year (2011–12), and

FIGURE 5: Percent of evaluators who reported engaging in specific activities on a daily or weekly basis



Source: ETW survey data, spring 2013.

they provided teachers with written and verbal feedback more frequently as well (See Figure 5).

Teachers confirmed this trend, reporting more frequent written and verbal feedback on their observations and lesson plans in fall 2012 than the previous year; roughly half of the teachers indicated they received written and verbal feedback more than three times in fall 2012 (See Figure 6). And more than half of the teachers in a Los Angeles Unified survey reported having increased instruction-focused interactions with their administrator.

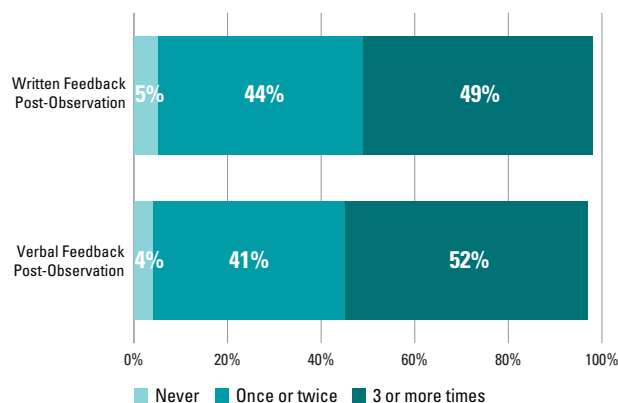
Many teachers described working in a professional culture where feedback is viewed as a support offered to maximize student learning, and 80 percent of all teachers surveyed agreed they receive valuable feedback on their instructional practice.²⁵ Nearly all teachers indicated they were comfortable with constructive feedback about their teaching (See Figure 7).

In addition, most teachers described the feedback they receive about their instruction as "objective" and "specific."²⁶ While not all teachers agreed that feedback on their teaching was as strong as it could be, one teacher shared her belief that it would improve: "The quality of the feedback is going to get better as the principal gets more familiar with the process. Next year will be even better."

Teachers reported high levels of encouragement to reflect on their teaching (96 percent) and to use evidence of student learning to inform their practice. In Los Angeles Unified, nearly three-quarters of teachers piloting the new evaluation system noted an increase in their reflection about instructional practice.²⁷

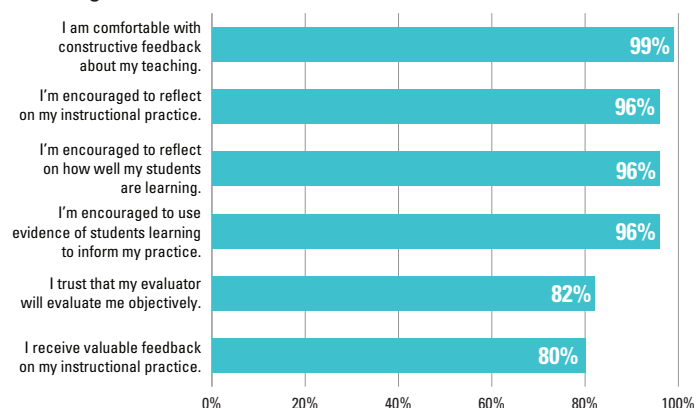
One of the most frequently reported challenges identified by educators in our study was that the evaluation process was very time-consuming, competing with other valuable

FIGURE 6: Frequency of written and verbal feedback reported by teachers



Source: ETW survey data, spring 2013.

FIGURE 7: Percent of teachers that agree or strongly agree with the following statements



Source: ETW survey data, spring 2013.

activities, such as collaborating with other teachers. It was not uncommon for teachers to take up to 10 hours to prepare a lesson plan as part of the formal observation cycle as they attempted to demonstrate nearly everything covered on the teaching framework rubric in one lesson.

Administrators also felt the time pressure, and fewer than 40 percent of them indicated they had sufficient time to coach, mentor, and support teachers. While several administrators shared with us that they were getting more efficient, 40 percent of the administrators surveyed reported it took them, on average, six hours or more to support a teacher through the evaluation process.²⁸ Master teachers in Lucia Mar Unified and dedicated leadership coaches in Alliance College-Ready Public Schools and Oakland Unified seemed to alleviate the challenges principals face with having so many varied responsibilities, and their approach may help other system leaders seeking to support principals who are feeling particularly overburdened.



SYSTEM SPOTLIGHT: PARTNERSHIPS TO UPLIFT COMMUNITIES SCHOOLS

Most individuals we spoke with described Partnerships to Uplift Communities as a learning organization that values reflection and revision. Teachers and administrators cited numerous opportunities for them to learn about, reflect on, and offer feedback about their entire Teacher Development System.

PUC central office administrators openly acknowledged that their initial launch of the TDS was overly focused on the mechanics of the evaluation process. However, they mediated that by developing “growth guides,” write-ups that describe a specific indicator of effective teaching and offer numerous resources and references to help teachers achieve the indicator.

Nearly 90 percent of PUC’s administrators indicated they provided input on the evaluation system. That feedback has had an impact. After principals and teachers made the case that the formal evaluation process was taking away opportunities for more frequent observations and occasions to provide feedback and coaching, PUC made another course correction. Rather than requiring one formal evaluation each semester, the CMO now requires one formal observation spring semester along with numerous informal observations throughout the year. PUC’s cofounder and CEO stated, “We were right to listen to the leaders and to take into consideration what they were saying.”

Affirming efforts made by the central office to ensure educators have a strong understanding of the evaluation criteria and more time to engage meaningfully with them, over 90 percent of PUC teachers and administrators reported feeling familiar with the evaluation criteria. Furthermore, in 2011–12, more than 90 percent of PUC teachers indicated someone had observed their teaching, and nearly 80 percent had observed another teacher’s instruction. One teacher echoed the sentiment of many PUC teachers, saying, “I feel like I’m being challenged to do better because of the TDS system. I like that people come into my classroom and give me feedback.” PUC administrators reported conducting observations and offering feedback to teachers more frequently than their peers across the other systems.²⁹

While the time demands for completing the evaluation cycle continue to be an issue, many school leaders appear to have streamlined the process, cutting the time it takes them to formally evaluate a teacher in half—from eight to 10 hours in the first year to four to five hours in the second. That efficiency appears to come in part from practice, explicit articulation of what counts as evidence for specific ratings, supportive technology, and extensive professional development and coaching opportunities for principals. One principal noted that she and her peers discussed the teaching framework and TDS in depth every week for an entire year, and that central office staff were always available to help principals improve on their evaluation and support efforts. Being more efficient with the formal process does not suggest that the amount of constructive feedback was curtailed. On the contrary, some principals indicated that they could spend more time engaging informally with teachers about their work when they were more practiced at documenting their formal interactions.

Over 95 percent of PUC teachers surveyed agreed they were encouraged to reflect on their instructional practice, to reflect on how well their students were learning, and to use evidence of student learning to inform their instructional practice. As one principal put it, “This system has allowed for leaders, coaches, and teachers to have much more critical conversation about practice. It hasn’t been about checking boxes off. It’s really afforded time for teachers and leaders to discuss lessons, practice, techniques, and strategies.”

The other time-related challenge we heard from many educators, and from school administrators in particular, was their struggle with competing priorities. However, in districts and CMOs where teacher evaluation and support was a top priority, principals felt clear about how they ought to focus their time.

Finding 5: Despite having some initial misgivings, most teachers are experiencing evaluation as a vehicle for improvement—not as a “gotcha”—and supports are becoming both more plentiful and better aligned with the evaluation criteria over time.

Nearly all educators told us that the goal of their teacher evaluation system is to improve student achievement by improving instruction, while just a few teachers stated the goal of their teacher evaluation system was to “eliminate incompetency.” Regardless of their take on the purpose, most teachers acknowledged they were more anxious about evaluation than they had been in the past.

For some teachers, the pressure came from what they described as increased expectations of both their teaching and students’ learning. For others, the strain came from knowing that their performance ratings could have an impact on their job security, compensation, and access to career advancement opportunities. Not all teachers felt overburdened by the evaluation expectations, however, acknowledging the prevalence of performance evaluations in other professional sectors and noting that teachers have been largely exempt from a rigorous evaluation process.

School leaders sought various ways to manage and mediate teacher apprehension to ensure supportive conditions for meaningful reflection. One principal, for instance, assured all teachers that the first year of the new teacher evaluation implementation would not have a negative impact on any teacher’s standing. Another principal reassured teachers that the anxiety they were experiencing was natural but would also dissipate as they gained experience within the new teacher evaluation process. Despite the challenges involved with increasing expectations for performance and improvement, most teachers had a positive view of the goals and the process overall. As one teacher noted, “My teaching improved. I’m always a willing learner, and this process took me to the next level.”

Most districts and CMOs that led with compensation instead of supports quickly learned they needed to reverse that and focus on developing early and ongoing professional development for both teachers and administrators. As one central office administrator noted, “We have teachers getting very specific feedback from the observations without knowing what to do with it, and principals do not have the systems to

“*I’m trying to get them to see that it’s about making improvements no matter where you are. This is evaluation that is meant to focus on teacher growth.*”
—Principal

address that.” Fortunately, there is evidence of an increasing number and variety of professional development opportunities aligned to the teacher evaluation rubrics in each of the districts and CMOs we have studied. In Los Angeles Unified, for example, while leaders initiated the teacher evaluation pilot without accompanying professional development, they adjusted the process to provide workshops addressing areas of growth most frequently identified by both teachers and evaluators.

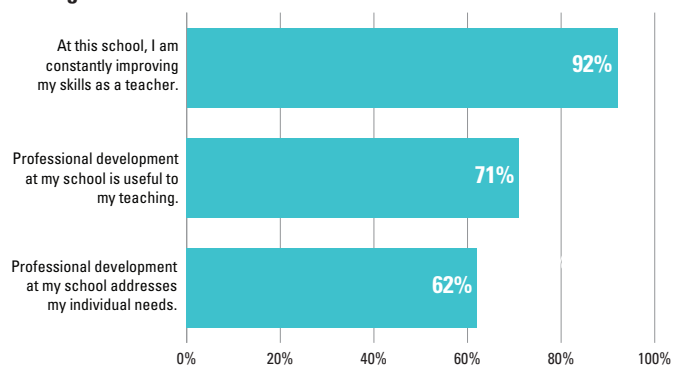
Some support opportunities are available anytime online. Teachers from various systems shared their excitement for rich descriptions of effective teaching and relevant videos illustrating research-based instructional strategies. Other professional development opportunities include expanding access to instructional coaches or master teachers with content-area expertise. Peer observation and feedback is more prominent in some systems than others and has improved, according to educators we interviewed, as teachers have developed more fluency with the language of effective teaching outlined in their evaluation frameworks and rubrics.

As administrators spent a greater share of their time on teacher development, they cited their efforts to link constructive criticism with specific supports, which included recommended workshops, articles, specific strategies, modeling, and videos. One administrator noted that teachers needing support beyond faculty-wide offerings at his school were offered tailored supports.

Close to two-thirds of the teachers indicated they accessed professional development tied to their individual needs,³⁰ and over 70 percent of teachers surveyed indicated that the professional development at their school was useful to their teaching.³¹ Finally, over 90 percent of the teachers surveyed noted they were constantly improving their teaching skills (See Figure 8).

Differentiating professional development to meet the needs of teachers across the performance spectrum is both critical and challenging. Newcomers to the teaching profession require intensive support in strategic areas, but highly accomplished veteran teachers need something entirely different. While some teachers lauded their site and system-wide capacity to offer professional development that met the varied needs of teachers and leaders, others

FIGURE 8: Percent of teachers that agree or strongly agree with the following statements



Source: ETW survey data, spring 2013.

complained that their needs were not being met. Despite this variation, almost every focus-group teacher we interviewed confirmed that they had professional development time built in to their regular workday; and others, though not all, received a stipend for weekly professional development meetings before or after school.

Finding 6: Training for both teachers and evaluators on the evaluation framework is absolutely critical. By working hard to ensure an extensive system-wide understanding of the criteria, and in most cases requiring evaluator certification, these school systems have created broadly shared understandings of and expectations for specific characteristics of good teaching.

One of the many things that sets the teacher evaluation processes in our study apart from typical ones is how deeply teachers and administrators study the teaching frameworks and apply them to their daily work. Teachers and administrators with expertise in the teacher evaluation criteria oriented their peers to the elements of effective teaching at the beginning of the school year, and continued to delve into these elements during regularly scheduled professional development sessions throughout the year. Student outcomes often informed what the teachers worked on. When students struggled to answer questions that went beyond factual recall, for example, teachers worked collaboratively on developing their questioning skills to better prepare their students for more challenging types of questions.

In addition, almost all of the systems require evaluators to be certified to conduct evaluations. To get certified, evaluators must demonstrate their knowledge of the evaluation criteria, their ability to gather relevant evidence, and their ability to accurately assess that evidence against the rubric. They do this by rating both a teacher's lesson plan and corresponding video-recorded



**SYSTEM SPOTLIGHT:
GREEN DOT PUBLIC SCHOOLS**

Green Dot Public Schools engages teachers in nearly all phases of its teacher evaluation planning and implementation. After acknowledging an early communication misstep regarding compensation under the new teacher evaluation system, Green Dot leaders shifted their messaging and resources to emphasize a focus on instructional improvement.

Green Dot leaders dubbed 2012–13 the “Year of Teacher Supports” and ushered in several changes to ensure teachers experienced tangible supports on every campus. For one, Green Dot system leaders hosted frequent sessions soliciting feedback from teachers and school leaders on the teacher evaluation and support implementation process. They also communicated updates in a weekly newsletter and launched various initiatives based on educators’ feedback. One such initiative was a pilot in 2012–13 for a new configuration of classroom observations known as the “many mini.” Instead of scheduling a few formal observations that seemed more like “dog and pony shows,” the many mini model calls for more frequent observations designed to capture more authentic classroom instruction.

Content-area coaches, hired to both provide more content-informed feedback to teachers and support the capacity of administrators to serve as instructional leaders, seemed to have a positive impact across schools. Nearly half of all Green Dot teachers surveyed indicated they had been observed by a coach, mentor, or lead teacher at least three times in 2011–12. One teacher described receiving a visit from her coach every other week to offer feedback, resources, and support in analyzing student data, noting, “It’s been very helpful getting constant feedback.”

In addition to coaching, many teachers receive administrator observations and feedback. Over 80 percent of the teachers surveyed indicated they were observed at least three times by a school administrator in 2011–12. Green Dot school administrators cited their efforts to link constructive criticism with specific supports, such as articles, strategies, videos, and modeling. These supports were evident in a one-on-one informal debrief between an administrator and teacher observed during a site visit.

Green Dot teachers also praised opportunities to observe and be observed by their peers, and survey results indicated that three out of four teachers observed someone else’s teaching and one in five reported being observed least three times by a peer in 2011–12. Most teachers indicated they received verbal and written feedback after an observation at least three times in 2011–12. Nearly 80 percent agreed that professional development at their school was useful to their teaching, and teachers in a focus group appreciated and would like to see more “efforts to differentiate supports for teachers based on their needs.” Teachers noted the value of “having the opportunity for authentic dialogue about how to improve their teaching,” and 65 percent of those surveyed agreed that the evaluation process was useful to them. However, study site visits suggested that the coordination between coaches and administrators could be improved, since it was not apparent that a teacher received consistent messages from coaches and administrators about how he or she might improve.

Despite this area of growth, when asked about whether they would ever want to go back to the way things were, teachers in focus groups generally agreed the new system was more meaningful. One teacher said, “It’s not perfect now, but I am getting a benefit and growing somewhat as a teacher.” Furthermore, teachers we spoke with indicated that the evaluation and support system was helping them to “constantly reflect on their practice and improve.”



SYSTEM SPOTLIGHT: LUCIA MAR UNIFIED SCHOOL DISTRICT

Lucia Mar Unified School District adopted TAP: The System for Teacher and Student Advancement in 2011–12 after spending a year researching and planning for its implementation as a way to support, develop, evaluate, and compensate teachers differently. What is remarkable about the teacher evaluation and support work in Lucia Mar is how deliberately teachers are supported to meet the increased expectations for effective instruction and how many respected, experienced educators describe their engagement with TAP as unequivocally “the best professional development” they have experienced.

In Lucia Mar, two master teachers have supported their fellow teachers full-time at each district TAP school. In order to qualify for this role, master teacher candidates earned “highly effective” ratings for their teaching, engaged in a rigorous selection process, and successfully completed TAP certification assessments. In addition to the structured supports they receive from the district’s master teacher coach, they can also access online supports to help them efficiently score and review evaluations.

Master teachers spend a significant amount of time facilitating meaningful weekly site-based professional development sessions organized by grade level (or content area). These sessions are driven by students’ needs and a clear description of effective teaching that includes 19 indicators across four domains. More than 80 percent of TAP teachers reported meeting in small groups to discuss their teaching, and nearly 85 percent also agreed that their school-based professional development was useful to their teaching—the highest ratings across the systems in our study. As one veteran teacher attests, “This is built-in professional development at our site with people that we trust and know well. If we’re struggling, [our master and mentor teachers] show us, they model [for us].” A coach adds, “With this model, you have fully released master teachers to support [the teachers], ... to keep following up to see if [they’re] implementing, [and to] help [them] take things to the next level.”

Principals, master teachers, and mentor teachers (full-time classroom teachers who take on additional responsibilities to support their colleagues) observe career teachers frequently and offer supports aligned to teachers’ individual professional growth goals as well as to school-wide goals. Of TAP teachers surveyed, 98 percent reported that

someone observed them teaching, and nearly 85 percent agreed they received valuable feedback. Not only did high percentages of TAP teachers report that they had had more than three observations in 2011–12 by their principal (45 percent), a coach (70 percent), and peers (30 percent), roughly half of the teachers also reported having experience with co-planning and/or co-teaching a lesson, and nearly three-quarters of the teachers had a lesson modeled for them.

Critical structures in the TAP model support a school-wide focus on instructional improvement. The school’s TAP leadership team, composed of the school administrators, master teachers, and mentor teachers, support the weekly professional development meetings. Together they review student data and determine the school goals, plan and conduct formal and informal teacher observations, and provide individual coaching guidance to teachers as they develop their growth plan. Three-quarters of the Lucia Mar TAP teachers surveyed indicated that professional development at their school meets their individual needs, as compared to less than 60 percent of the teachers surveyed in non-TAP schools. Furthermore, more than 80 percent of the TAP teachers in Lucia Mar agree that “the evaluation process is useful” versus 65 percent of non-TAP teachers. One TAP administrator noted, “In the past [teacher evaluation] was so superficial and subjective. ... You were competent or incompetent—there was no in between. This process allows teachers to hear feedback about their teaching and receive quality suggestions.” Lucia Mar’s TAP leaders have become insightful coaches intently focused on supporting teachers’ efforts to improve their teaching.

Based on the 2012–13 external evaluation report for the TAP schools in Lucia Mar, the changes in teaching practice are associated with improvements in student performance. What’s more, the evaluators report that the TAP system in Lucia Mar has helped to close the achievement gap between low-income and more affluent student populations, as observable improvements were made in the percentage of students scoring proficient or advanced in English language arts and math in TAP schools compared to matched non-TAP schools.³² Furthermore, elementary English learners in TAP schools outperformed their non-TAP peers on the ELA and math CSTs.³³

“*[The rubric] gives me a solid foundation on what good instruction is. It gives me a focus for my conversation with teachers. I’ve become a better instructional leader through this process.*”
—Principal

classroom lesson against the instructional framework. If their ratings match the ratings established by experts, they earn certified status and can officially evaluate teachers. Not all candidates are competent at marshaling evidence to appropriate indicators or accurately assessing the evidence against the criteria. In fact, in many of the systems we studied, fewer than 50 percent of certification candidates passed the certification assessment on their first try. When a candidate does not pass the certification assessment, an evaluation coach typically offers the candidate

additional support and practice—often by partnering with them to rate formal observations—until he or she is able to get certified.

Despite various supports, only roughly half of all evaluators surveyed felt adequately prepared to evaluate teachers. That figure varied considerably by system, with as few as 39 percent of evaluators feeling sufficiently prepared in one system and as many as 86 percent feeling sufficiently prepared in another. Overall, approximately 70 percent of evaluators indicated that “evaluators are consistent in how they evaluate teachers.” In most systems there is room to grow to ensure that everyone holds the same understanding and expectations of effective teaching. To address this need, professional development focused on effective teaching and rigorous evaluation certification requirements contribute to a common set of understandings of and expectations for specific effective instructional practices.

Finding 7: Virtually every school system made significant changes to its teacher evaluation and support process along the way; indeed, the willingness to do so when problems occurred helped build trust and confidence.

Across each of the systems both teachers and evaluators had opportunities to provide feedback on both the evaluation criteria and process. Formal opportunities for feedback included focus groups, surveys, and sessions to revise the instructional rubrics and other aspects of the teacher evaluation process. However, educators' perceptions of how much efficacy they have on the process varied by system and by role. While 52 percent of teachers and 73 percent of evaluators overall said they had input into their current evaluation process/system (See Figure 9), levels of agreement among educators differed widely within and across the seven systems.³⁴ In some systems, much more could be done to engage educators in a continual improvement process.

We encountered teachers who expressed frustration that the feedback they had offered was not acknowledged. Nevertheless, teachers generally affirmed consistent efforts on behalf of system leaders to gain valued input, and they identified numerous changes resulting from that feedback, notably revisions to the framework, adjustments to observation requirements, and increases in the professional development supports available to teachers.

In some systems, the teachers' union has played a significant role in ensuring teacher input in and transparency of the evaluation process.³⁵ As a Green Dot teacher confirmed, "The union plays a huge part in this too. Every step of the way, management has reached out to the union. We've heard presenters talk about it, we have offered feedback through surveys and focus groups; they offer webinars to explain. This was voted on by teachers—we [teachers] decided to take it on." According to several CMO leaders, members

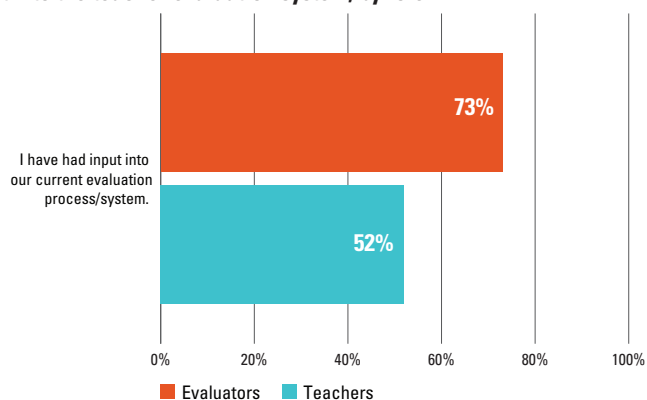
and leaders of Green Dot's teacher union contributed significantly to the initial development of TCRP's teaching and learning framework, and offered important subsequent recommendations during each cycle of revision.

Besides the general discrepancy between teacher and evaluator perspectives on feedback, each of the systems struggled to address the issues and concerns raised by different groups of educators. Specifically, many teachers in non-tested grades and content areas³⁶ voiced concerns about being evaluated fairly, since the student-growth component of their evaluation was typically based on school-wide measures rather than measures specific to students they taught. In addition, many first-year teachers felt overwhelmed by simultaneous demands placed on them by their Beginning Teacher Support and Assessment program and the teacher evaluation system.

System leaders prioritized which issues they would address at any given time, but often with limited resources and capacity to do so. At times these leaders also struggled to clearly communicate which issues they would address and how, further frustrating affected educators.

Notwithstanding varied perceptions among educators about how much influence they have had on the teacher evaluation system, leaders in each system have solicited and taken seriously feedback from teachers, principals, other evaluators, and coaches; and they have made substantive changes to the teacher evaluation process based on that feedback. Suggestions that emerged through dedicated and structured feedback from teachers, principals, coaches, and other key stakeholders in the evaluation system often led to system improvements. As illustrated by several examples, an initial design rarely ends up becoming a permanent design. Alliance College-Ready Public Schools introduced teachers as trained evaluators system-wide after noting the positive impact teacher evaluators were having on the quality of teacher feedback and the ability to provide talented teachers with career growth opportunities at one school site. When Lucia Mar Unified system leaders faced opposition from the teachers' union over the implementation of TAP, they decided to require two observers for every formal observation in order to mitigate the opposition and build trust in the process. Although time-consuming, this policy helped expand the instructional leadership capacity of mentor teachers, master teachers, and principals, and helped observers calibrate their ratings. In Oakland Unified, instructional coaches responded to teachers' desires to have concrete examples of effective instructional practices by capturing them on video and deconstructing them with teachers to increase their use. Overall, these examples emphasize the importance of engaging key stakeholders in a constant cycle of improvement.

Figure 9: Percentage of educators who agree or strongly agree they had input into the teacher evaluation system, by role



Source: ETW survey data, spring 2013.



SYSTEM SPOTLIGHT: LOS ANGELES UNIFIED SCHOOL DISTRICT

In 2009 the Los Angeles Unified School Board established the Teacher Effectiveness Task Force, which laid the groundwork for the teacher evaluation pilot in 2011 known as the Teacher Growth and Development Cycle. Over 1,000 educators contributed to the hefty set of 63 teaching standards constituting part of the district's new Teaching and Learning Framework, but the pilot launched amid a heated political fight between the district and teachers' union, United Teacher Los Angeles. Both sides differ in how to address a court order from 2011 to enforce the Stull Act's provision that student achievement results be included as part of the formal evaluation process for teachers. While an agreement was reached in November 2012 and ratified in January 2013, UTLA asserts that Los Angeles Unified arbitrarily imposed the TGDC.

According to district administrators, the primary focus of TGDC is to improve instruction for the 600,000-plus students in the district by getting teachers the support they need and deserve as professionals. The primary purpose of the pilot was to test and adapt the evaluation tools toward that end.

The first-year pilot involved over 70 principals, roughly 400 teachers, and central office staff that included 12 Teaching and Learning Coordinators. TLCs provided ground-level support to pilot principals and teachers, working side by side with them through the evaluation cycle. Together they reviewed written lesson plans, conducted pre-observation conferences and classroom observations with principals, and rated teacher performance according to the Teaching and Learning Framework rubric. They also helped manage other aspects of TGDC, troubleshooting issues as they arose and making design and process improvements along the way.

Los Angeles Unified also contracted services from WestEd (a nonprofit research, development, and service agency) and the University of Southern California to administer four surveys and conduct 60 focus groups. Despite these extensive efforts to gather feedback and use it to inform adaptations with the TGDC tools, some teachers perceived that little if any of their feedback would be put to use, and district administrators acknowledged that people were not always aware how deliberately they were collecting and responding to feedback about the tools and process.

Many teachers voiced their concerns about what they considered to be a "clunky" and "glitchy" pilot implementation that demanded considerable amounts of time. As one teacher put it, "We volunteered to do this and we're having difficulty. How are teachers who are resistant going to face this?" Despite the challenges, some educators described how the pilot has improved on the evaluation process they had experienced in the past. One teacher said, "We didn't have this clarity before. This gives clarity to principals and evaluators about what they should be looking at. ... We are on the same page, and when the principal comes in, we know what she is looking for."

TLCs and district administrators made several major improvements in the TGDC pilot between 2011–12 and 2012–13. One significant change was reducing the number of focus elements required for the classroom observation from 61 to 15 and revising elements that required more clarity. As one principal noted, "There were parts of the framework where the language was off. We made suggestions and we did see changes." TLCs also redesigned the lesson template to make it shorter, less complex, and less redundant. And they trimmed the teacher self-assessment and Individual Growth Plan protocols as well.

Similar to other systems, Los Angeles Unified offered teacher training and supports aligned with the rubric. In March 2013, for example, 300 educators attended a professional development conference designed to support elements of the Teaching and Learning Framework most frequently identified by teachers in their IGPs. And according to multiple stakeholders, teachers were becoming more proactive about their professional learning.

Simultaneously, efforts to improve the technology platform supporting TGDC made it easier and more reliable to use. Pilot volunteers who described frustrating bugs with the technology platform—losing work, getting stuck, or ending up with the wrong observation notes—acknowledged that efforts to streamline the online process for teachers and principals made a big difference in their user experience. Overall, TGDC experienced a significant transformation from its first-year pilot to the second-year pilot, and it continues to evolve with implementation amid a politically charged environment.

Recommendations

Based on what we have learned so far, statewide policy-makers and district leaders can take several actions to overhaul teacher evaluation and support systems across the state.



WHAT STATE DECISION AND POLICYMAKERS CAN DO

1. **Strengthen state guidelines and requirements for teacher evaluation.** California lawmakers should adopt a statewide framework to guide districts' local adoption of improved teacher evaluation systems. This framework should direct districts to align what students, teachers, and school leaders are expected to do. Specifically, teachers should be evaluated against a set of standards that reflect and reinforce the curriculum standards for students. Likewise, districts should evaluate school leaders

against a set of standards that aligns with what teachers are expected to know and be able to do. In addition, the framework should also direct districts to conduct more frequent evaluations; should incorporate multiple measures of teaching performance, including but not limited to objective measures of student achievement growth and frequent observations; and should make evaluations routine for all teachers, regardless of their years of experience. Legislators should specify that this evaluation system be coupled with a system of individualized and ongoing professional development that provides teachers with the strategies, mentoring, and support they need to grow as professionals. When developing this framework, state leaders can draw from many promising models already being implemented in California districts and charter management organizations, including those highlighted in this report.

2. **Remove barriers to staffing decisions that are in the best interests of students, with particular emphasis on high-need students.** The state should eliminate its seniority-based layoff requirement whereby the most recently hired teachers, regardless of their effectiveness, are the first to be laid off when a school district faces financial uncertainty.³⁷ Lawmakers should grant school and district leaders the right under state law to protect effective teachers from layoffs, to hire teachers based on students' needs, and to decline the forced placement of ineffective staff. If a district implements an evaluation process that differentiates performance among teachers, then school and district leaders must be able to act on it.
3. **Leverage funding sources to support local teacher evaluation work.** Ongoing costs for teacher evaluation should come from district budgets. In California, districts' newfound flexibility and increasing revenues (thanks to the Local Control Funding Formula, Proposition 30, and strong economic growth) make this more feasible than it was during the Great Recession. That said, the state and school districts should continue to leverage existing federal programs, including the School Improvement Grant program and Teacher Incentive Fund, to fund this work while also seeking creative ways to leverage Title 1 Program Improvement set-asides to support this work. We also suggest the state identify short-term streams of state funding to jump-start this work and support early implementation efforts, much as it has done with the Common Core or as the U.S. Department of Education has done with Teacher Incentive Fund awards.
4. **State policy should require districts to monitor whether low-income students and students of color have equitable access to effective teachers.** Data on teacher effectiveness should be aggregated at the school, district, and state-levels, but should not be reported by individual teacher.
5. **Develop programs and policies to place and retain the best teachers in our highest-need schools.** In addition to establishing consistently held expectations for effective teaching at each phase of teacher development (e.g., recruitment, preparation, certification, induction, evaluation and on-going professional development support), the state should also develop programs and policies that ensure students in our highest-need schools have equitable access to effective teachers.



WHAT DISTRICT AND SCHOOL LEADERS CAN DO

6. **Make teacher evaluation and support work a top priority.** If district leaders do not do this, teacher leaders, school leaders, and/or community leaders need to press board members to make it one. Otherwise teacher evaluation and support will not be afforded the time, attention, or resources necessary to overcome inevitable challenges. District leaders should commit substantial resources, including dedicated staff, to the effort to transform teacher evaluation and support systems. Because a teaching evaluation system cannot function well as an isolated initiative, leaders must intentionally align it with hiring and induction processes, professional development and support, leadership coaching, and student supports.
7. **Ensure all key stakeholders are engaged.** Teachers, administrators, central office staff, community members, and labor union representatives must be part of the effort to implement a multi-measure teacher evaluation system. Leaving out any of these stakeholders could jeopardize efforts to build the trust necessary to support and sustain a collective effort. This is important not just at the beginning, in the design process, but all along the way; formal structures for on-going teacher input and feedback are critical to every stage of developing and implementing an effective teacher evaluation and support system.
8. **Adopt or create a framework of quality teaching with clear indicators that are aligned to standards for student learning.** There is no need to start from scratch, and district and school leaders must be careful to avoid unwieldy frameworks; if they are too cumbersome they will defeat the purpose of having a manageable and measurable description of high-quality teaching. Leaders should be sure the teaching frameworks are aligned with the standards for student learning, specifically Common Core State Standards, Next Generation Science Standards, and English Language Development Standards.
9. **Establish multiple measures of teaching quality and make sure they are valid and reliable.** No single measure can perfectly capture a teacher's craft, and most teachers welcome having a multifaceted review of their work. That said, when teachers and school leaders question measures by which they are evaluated, the entire system is vulnerable. Maintaining high rates of trust among educators requires frequent and strategic dialogue, collective knowledge-building, and intensive training.

- 10. Train people thoroughly on both the evaluation indicators and process.** School leaders should anticipate the need for a heavy initial investment to acculturate teachers, principals, instructional coaches, and central office staff to a new set of measures for teaching. They should also plan for ongoing, annual professional development to ensure teachers and administrators new to the system are adequately introduced to the organization's vision and expectations, and are clear about the process and supports designed to help them improve.
- 11. Establish a robust certification process for evaluators and ensure they are empowered to meaningfully differentiate teacher effectiveness.** To establish trust among teachers that evaluators are capable of evaluating their teaching and students' learning, every evaluator ought to be certified annually through a rigorous process. Certification can help ensure that evaluators are calibrated to the standards of teaching effectiveness and competent with the process. Individuals not meeting the certification criteria should be matched with certified evaluators and/or coaches who can work with them until they are certified. Individually and collectively, evaluators need to be able to meaningfully differentiate teaching performance.
- 12. Develop a detailed plan for all stages of implementation, including a pilot phase.** Determine how often teachers will be observed, what each individual's role(s) will be in every phase of implementation, and what resources, including dedicated personnel, are necessary to carry out the plan. The plan should be flexible and allow for adjustments. Be prepared to call upon experts in the field to bolster design and implementation efforts and take advantage of partnerships and/or networks to help develop and sustain the work.
- 13. Identify structures and resources for instructional support in advance.** Align instructional supports to evaluation frameworks so that teachers feel supported, not criticized. Ensure educators have opportunities to develop a shared language to describe and understand high-quality instruction. These supports will allow teachers to reflect, self-evaluate, and establish personal goals. Consider which informational and human resources (e.g., videos, artifacts, instructional coaches, collaborations, professional development) will be necessary to address a full range of identified growth areas.
- 14. Pay particular attention to principals.** Be sure to prioritize instructional effectiveness as a central part of the principal's job, and provide principals

with the supports they need to be successful. These supports include developing strategies for effectively managing their time and improving their coaching skills, adequately preparing them to develop and lead professional development, making sure other evaluators are in place to share observation responsibilities, and aligning principal and teacher evaluations and supports.³⁸ Consider also reorganizing the school and/or central office staffing to free up more principal time for instructional leadership.

15. Accept that the system will take time to fine-tune.

Acknowledge up front that the system will not be perfect in its first year, and prepare to make adjustments along the way. Given those likely changes, it also makes sense to keep initial stakes low, which will also help build trust in the system as people are getting familiar with it

Conclusion

When Education Trust staff return from visiting unusually effective schools serving low-income students and students of color, their most indelible memories aren't of the supports these schools provide for students—though those are often wonderfully robust—but of the supports they provide for teachers. "At that school," they often say, "even teachers who would be merely ordinary anyplace else become superstars."

Across California, teachers crave honest feedback and the support they need to become ever more effective. And we know from long experience that if more of them got that feedback and support our children would benefit hugely. High student performance that is unusual today could actually become commonplace.

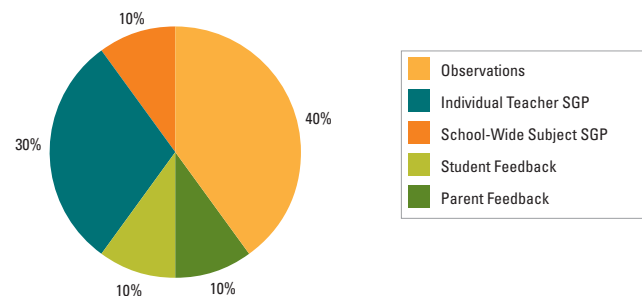
Our state has made important investments on one side of the equation, by putting generous resources into adopting a framework for teaching and into teacher development programs such as Beginning Teacher Support and Assessment and Peer Assistance and Review. But maintaining an evaluation framework that gives nearly every teacher the same performance rating seriously undercuts the effectiveness of these investments.

If we are serious about addressing inequities and giving our children a world-class education, this has to change. And now is the perfect time: We've adopted a new set of standards for our children—the Common Core State Standards in English language arts and mathematics, Next Generation Science Standards, and English Language Development standards. Now, we need to bring our standards for teaching into line.

We hope the findings and recommendations in this report will be helpful to state and district leaders as they contemplate how best to do that.

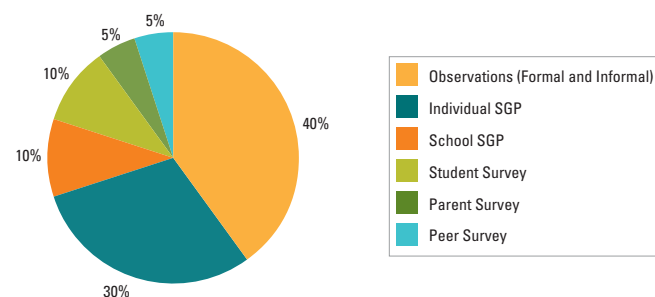
Appendix A: Evaluation Metrics by System

FIGURE 10: Alliance – CST-tested Subjects, 2012–2013



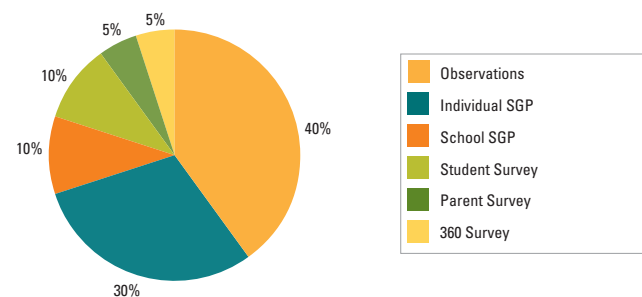
Source: Alliance artifacts and interview data.

Figure 11: Aspire – CST-tested Subjects, 2012–2013



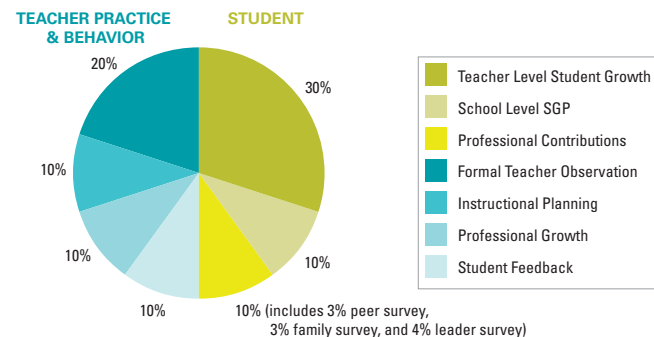
Source: Aspire artifacts and interview data.

FIGURE 12: Green Dot – CST-tested Subjects, 2012–2013



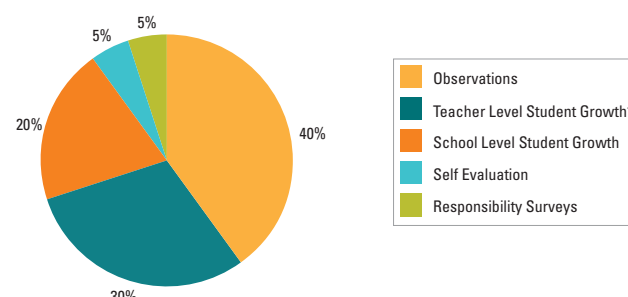
Source: Green Dot artifacts and interview data.

FIGURE 13: PUC – CST-tested Subjects, 2012–2013



Source: PUC artifacts and interview data.

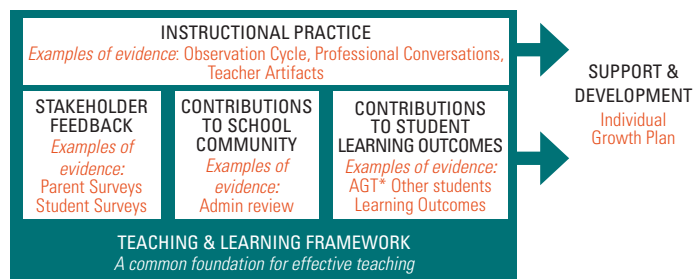
FIGURE 14: Lucia Mar – CST-tested Subjects, 2012–2013



*Value-added model is used by TAP to measure student gains.

Source: Lucia Mar artifacts and interview data.

FIGURE 15: Los Angeles Unified's Teaching and Learning Framework for Pilot Years 2011–12 & 2012–13



*Academic Growth over Time is a way to measure student gains.

Source: Los Angeles artifacts and interview data.

FIGURE 16: Oakland Unified's Teaching Evaluation Project Measurement Sources, 2012–2013

Domains	Lesson Plan and Pre Conference	Classroom Observations	Reflective and Post Conference	Student Feedback	Peer and Supervisor Feedback
DATA DRIVEN PLANNING AND ASSESSMENT	X		X		
CLASSROOM LEARNING ENVIRONMENT		X		X	X
INSTRUCTION		X		X	
PROFESSIONAL RESPONSIBILITIES		X			X
PARTNERSHIPS, FAMILY, AND COMMUNITY					X

Source: Oakland Unified artifacts and interview data.

NOTES

1. While many districts and charter management organizations rely exclusively on principals or other administrators to evaluate teachers, some systems allow for educators in other roles to rate teachers' classroom instruction. Lucia Mar Unified School District, for example, relies on master and mentor teachers to observe and evaluate teachers' classroom performance.
2. See, for example, Ray Chetty, John Friedman, and Jonah Rockoff. "Measuring the Impacts of Teachers II: Teacher Value-added and Student Outcomes in Adulthood," NBER working paper 19424 (September 2013). Also, H. Johnson, and R. Sengupta, "Closing the Gap: Meeting California's Need for College Graduates." (San Francisco, CA: Public Policy Institute of California, 2009).
3. Carrie Hahnel and Orville Jackson, "Learning Denied: The Case for Equitable Access to Effective Teaching in California's Largest School District" (Oakland, CA: The Education Trust–West, 2012). Also, Jeffrey Max and Steven Glazerman, "Do Disadvantaged Students Get Less Effective Teaching?: Key Findings from Recent Institute of Education Science Studies" (Washington, D.C.: Institute of Education Studies, NCEE Contract number: ED-04-CO-0112/0007, January 2014). Also, Michael Kent, Dennis Li, and Chris Frank. "Teacher Evaluation, Development, and Dismissal in California Part 1: The Measurement Gap: Investigating the Accuracy of Teacher Performance Evaluations," (Stanford, CA: Stanford Institute for Economic Policy Research, May 2013).
4. The California Standards for the Teaching Profession were first developed in 1997 and updated in 2009. For details see: <http://www.ctc.ca.gov/educator-prep/CSTP/CSTP.pdf>. and <http://www.ctc.ca.gov/educator-prep/standards/CSTP-2009.pdf>.
5. California Department of Education, 2013, "California NAEP Results," available at: <http://www.cde.ca.gov/tg/tg/nr/caresults.asp>.
6. Remediation rates for California State University first-time freshman students was over 40 percent, according to the Analytic Studies Division compiling data from 23 campuses of the CSU. Data available at: http://www.asd.calstate.edu/performance/combo/2013/Combo_Prof_Sys_Final_Fall2013.htm.
7. National Council on Teacher Quality, *2013 State Teacher Policy Yearbook Report* (Washington, D.C.: NCTQ, 2014).
8. A school system refers to either a district or a charter management organization.
9. Roughly half the respondents were high school teachers, under a third were middle school teachers, and nearly a quarter were elementary teachers. More than half of the teachers (60%) had fewer than six years of teaching experience, and 15 percent had been teaching at their current school for more than five years.
10. Oakland Unified, for example, incorporated a new teacher evaluation system for schools that won a federal School Improvement Grant. Lucia Mar Unified School District, on the other hand, applied for a Teacher Incentive Fund grant that funded several of their elementary and middle schools to implement TAP: The System for Teacher and Student Advancement. After Los Angeles Unified school board members passed a resolution in April 2009 establishing a Teacher Effectiveness Task Force, the district piloted its teacher evaluation system in 2011–12 with the help of TIF and SIG funds. The College Ready Promise, formed by four CMOs and funded by the Bill & Melinda Gates Foundation, committed to implement teacher and administrator effectiveness reforms.
11. Notably, teachers with fewer years of experience were more likely to agree that the evaluation process was useful to them (85 percent for first-year teachers) than teachers who had more than 10 years of experience (69 percent).
12. Similarly, teachers from systems with a union were somewhat less likely to agree (<90%) that more effective teachers should be offered increased leadership opportunities as compared with teachers from non-union systems (>90%).
13. The College Ready Promise rubric defines indicators for 15–16 standards across four domains, and this is used by four charter management organizations in TCRP: Alliance College-Ready Public Schools, Aspire Public Schools, Green Dot Public Schools, and Partnerships to Uplift Communities Schools. There are a few minor differences among each CMO's version of the college-ready teaching framework. Oakland Unified School District School Improvement Grant schools adopted the TCRP framework for their evaluation framework. The TAP: The System for Teacher and Student Advancement rubric used in Lucia Mar Unified defines 19 criteria across four domains. The Los Angeles Unified Teaching and Learning Framework articulates 61 elements across five standards.
14. Charlotte Danielson's Framework for Teaching divides teaching into 22 components across four domains. For details see: <http://danielsongroup.org/framework/>.
15. In Los Angeles Unified, 76 percent of teachers and 82 percent of observers completing the survey on the 2011–12 Initial Implementation Phase reported that their training gave them "full understanding" of the district's Teaching and Learning Framework, with a number of IIP participants suggesting (through open-ended responses) the need for deeper training, especially for teachers, on the framework. For more details see: <http://lausd.schoolwires.net/cms/lib08/CA01000043/Centricity/Domain/433/IIP%20year%20end%20feedback%20report.pdf>.
16. On the Los Angeles Unified–conducted surveys for the 2011–12 Initial Implementation Phase, 96 percent of observers and 85 percent of teachers agreed that "evaluations were based on a set of standards that reflect my definition of effective teaching." And over 80 percent of educators agreed that "the teaching and learning framework provides an appropriate description of effective teaching for the subject and grade I teach."
17. Alliance College-Ready Public Schools, Aspire Public Schools, Green Dot Public Schools, and Partnerships to Uplift Communities.
18. Eighty-two percent of teachers in the Education Trust—West's 2013 spring survey agreed that they "trusted their evaluator to evaluate them objectively," and 83 percent of Los Angeles Unified's 2013 spring survey teacher respondents "felt the observation score they received was fair" while 85 percent felt that the "evidence collected was objective" and 80 percent felt that "the evidence collected was thorough." For more details: <http://talentmanagement.lausd.net/sites/talentmanagement.lausd.net/files/Docs/EGDC/Teachers/EGDC%202012-13%20midyear%20feedback%20report.pdf>.
19. In California, not all students are tested with state-mandated standardized tests. For example, students in kindergarten and first grade are not tested at all. And not all content areas require state-mandated standardized tests. For example, art, music, world language, and physical education courses do not have to mandatory statewide tests.
20. See Ronald Ferguson, "Can Student Surveys Measure Teaching Quality?" *Phi Delta Kappan*, v94, n3 (November 2012), 24–28, <http://tripodproject.org/wp-content/uploads/files-common/Kappan-11-2012.pdf>
21. Each student's growth percentile is calculated by comparing the change in his or her California Standards Tests performance to all other students in The College Ready Promise and Los Angeles Unified who had the same CST achievement results in previous years (the student's "academic peers"). A teacher's SGP rating is the average SGP for students in his or her class that met attendance criteria for the year.
22. Similarly to the teacher SGP rating, a school-wide SGP rating can be calculated.
23. In Lucia Mar, for example, two full-time master teachers spend most of their time observing and coaching teachers at each TAP school. In other systems, such as Aspire and Green Dot, principals are most often a teacher's primary observer.
24. See Thomas J. Kane and Douglas O. Staiger, "Gathering Feedback for Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains (Seattle, Wa.)" The Bill & Melinda Gates Foundation: January 2012. A key finding from this Measuring Effective Teaching study is that multiple teacher observations conducted by multiple observers yielded more reliable evaluations than those conducted with limited observations and one observer.
25. The spring 2013 Los Angeles Unified survey posed questions to teachers and observers about various aspects of the Educator Growth and Development Cycle process, including how meaningful aspects of the self-assessment, lesson development template, pre-observation conference, observation, and post-observation conference were. Of the teachers surveyed, 80 percent said they had meaningful discussion with their administrators about the self-assessment, lesson plan, and their classroom instruction, and 70 percent said they received useful feedback about improving their lesson plan.
26. Los Angeles Unified survey data on the 2011–12 Initial Implementation Plan revealed roughly 60 percent of teachers "gained a better understanding of their individual needs as a teacher" and reported an increase in receiving more detailed information about their individual teaching practice.
27. In the Los Angeles Unified survey on the 2011–12 Initial Implementation Plan, 74 percent of teachers reported a "great" or "moderate" increase in reflection about instructional practice.
28. Data from the Los Angeles Unified spring 2013 survey yielded more specific information on how long various components of the evaluation process took teachers and evaluators to complete and whether they felt that amount of time was an appropriate amount of time or took too long. For example, nearly half of the observers (47%) spent between three and five hours tagging evidence from the observation notes and aligning it with the instructional rubric; and 85 percent of observers felt that tagging evidence took too long.
29. In fall 2012, for example, nearly 70 percent of PUC administrators reported providing verbal feedback to teachers daily or weekly, as compared to 40 percent of administrators across all six systems surveyed. Likewise, 40 percent of PUC administrators reported conducting observations on a daily or weekly basis as compared to 30 percent of administrators across all six systems surveyed.
30. With as few as 50 percent agreeing to that in one system and as many as 80 percent agreeing in another.
31. With as few as 62 percent agreeing to that in one system and as many as 80 percent agreeing in another.
32. James L. Gentilucci and James J. Brescia, "Lucia Mar Unified School District: Evaluation Report: TAP Instructionally Focused Accountability System." California Polytechnic State University, November 2013. <http://www.luciamarschools.org/wp-content/uploads/2013/12/LMUSD-TIF-Evaluation-Report-2013.pdf>.
33. English learners at TAP elementary schools who tested as "proficient" in ELA increased from 33.3 to 33 percent, as compared to a 1 percent decline for ELs in non-TAP elementary schools, (from 28.3 to 28 percent proficient). Likewise, ELs at TAP elementary schools outperformed their non-TAP peers in mathematics, growing from 46 to 54 percent proficient versus ELs at non-TAP schools who grew from 36 to 43 percent proficient.
34. Teacher agreement ranged from 32 to 70 percent, and evaluator agreement ranged from 33 to 86 percent.
35. Four systems have teachers' associations that impact the teacher evaluation and support system: Green Dot Public Schools, Los Angeles Unified, Lucia Mar Unified, and Oakland Unified. In two of these districts, the relationship between the teachers' association and the district administration has been under significant strain throughout the development, piloting, and early implementation phases of the teacher evaluation system. Specifically, United Teachers of Los Angeles and Los Angeles Unified have clashed over ways to address a court order resulting from a lawsuit brought against Los Angeles Unified in 2011 to enforce the Stull Act's provision that student achievement results be included as part of the formal evaluation process for teachers. In Lucia Mar Unified, the teacher's association battled the school district over contract negotiations in 2012, positing that the TAP program had to be renegotiated. An agreement was reached in August 2012 allowing for individual TAP schools to determine through a teacher vote whether they would stay engaged in TAP or not. In summer of 2013, one school opted out. In spring 2014, another school opted out. For the 2014–2015 school year, five sites out of the original seven schools remain in TAP. In Oakland Unified, the union has been active in a process to determine a district-wide teacher evaluation system in a parallel effort to the SIG schools, but has not vigorously challenged the SIG model. At Green Dot Public Schools, teachers characterized the relationship between many teachers and management as "open" and "productive."
36. This group includes K–2 teachers, visual and performing arts teachers, some science and social studies teachers, and special education teachers.
37. Current law requires that districts lay off teachers in inverse seniority order, but it provides some exceptions for deviating from seniority to protect specialized teachers or to achieve equal protection of the laws. For more detail see EdCode Section 45117: <http://law.onecle.com/california/education/45117.html>.
38. See Paul Bambrick-Santoyo, Doug Lemov, and Brett Peiser. *Leverage Leadership: A Practical Guide to Building Exceptional Schools* (San Francisco, CA: Jossey-Bass, 2012.)

ACKNOWLEDGEMENTS

We offer our appreciation to the Bill & Melinda Gates Foundation and the Stuart Foundation for their generous support which made this report possible. We also offer our deep gratitude to the individuals in school systems who shared their time, insights, and experiences with us. In addition, the entire Education Trust-West staff, along with project consultant Patrick Lee, provided invaluable support in conducting this research and preparing this report.



The Education Trust—West

1814 Franklin St., Suite 220, Oakland, Calif. 94612

T 510/465-6444 • F 510/465-0589

www.edtrustwest.org