



Aligned by Design

How Teacher Compensation Reform Can Support
and Reinforce Other Educational Reforms

Craig Jerald July 2009



Aligned by Design

How Teacher Compensation Reform Can Support
and Reinforce Other Educational Reforms

Craig Jerald July 2009

Introduction and summary

Over the past 10 years, long-standing approaches to compensating teachers in primary schools, middle schools, and high schools have come under increasing criticism. The so-called single-salary schedule, which emerged in the 1920s as a way to make teachers' pay less arbitrary and more equitable, seems highly inefficient in an era where education policy seeks to improve student outcomes and education systems must aggressively compete with other sectors for talent.¹ The single-salary schedule used in most places:

- Rewards only experience and graduate education courses, which have been found to be weakly or even negatively associated with student achievement
- Provides administrators no flexibility to respond to market forces
- Offers teachers no financial incentive to improve their instructional expertise and effectiveness
- Forces even the most effective teachers to wait many years to reach the higher rungs of the schedule, undermining recruitment and retention of talented college graduates.²

As a result, policymakers at the national, state, and local levels have proposed or enacted a variety of policies to better differentiate teacher pay. Indeed, nationwide, there is more experimentation with teacher compensation reform than any time since the *A Nation at Risk* report spurred a slew of “merit pay” and “career ladder” initiatives in the mid-to-late 1980s.³ To address problems with the traditional salary schedule, current initiatives typically focus on one or more alternative ways to differentiate teacher compensation:

- Pay for performance based on outputs—the performance of teachers’ students.
- Pay for skills and knowledge based on inputs—the value of teachers’ varied abilities
- Pay for hard-to-staff subjects or locations based on local labor market conditions
- Pay for additional roles and responsibilities based on higher workloads.⁴

Based on recent initiatives in a number of states and localities, along with the mostly unsuccessful attempts at reform 20 years ago after the publication of *A Nation at Risk*, experts now point to a number of lessons learned that should be examined when designing compensation reforms in order to make such reforms workable and sustainable over a long enough period to have a positive impact.⁵

One key lesson learned is the relationship between teacher compensation and other strategies for recruiting, developing, and retaining a high-quality teaching work force.

Reviewing studies of several local performance pay programs, Herbert Heneman, Anthony Milanowski, and Steven Kimball, researchers with the Consortium for Policy Research in Education, found that performance pay was nearly always implemented as a “standalone” reform not linked to schools districts’ broader improvement plans or to other human resources policies. That lack of alignment, they conclude, hinders the sustainability and impact of the performance pay initiatives in the districts examined in the studies.⁶

Allan Odden, co-director of CPRE and CPRE’s new project on Strategic Management of Human Capital in education, offers an example:

We have worked in districts that have developed knowledge- and skills-based pay structures, using a performance evaluation of teachers with a specific set of teaching standards and scoring rubrics, but for the first ten years did not align the professional development with the teaching practice embedded in the new evaluation system.⁷

In other words, the districts began paying teachers to develop certain kinds of specific knowledge and skills—such as acquiring new competencies in curriculum development or improving classroom instructional skills described in formal standards for evaluating good teaching—then neglected to ensure that district-funded professional development activities were focused on the same knowledge and skills.

Based on these school districts’ experiences, Odden and his colleagues conclude that “a revised teacher and principal pay structure by itself will have a modest effect if the other parts of the human resource management system are not realigned.”⁸

Other experts on compensation reform have begun to echo such conclusions. “I don’t think performance pay is enough on its own,” says the Urban Institute’s Dan Goldhaber, who has conducted extensive research on the topic. “If all you do is plunk down a pay-for-performance model and it’s not implemented well and you don’t have data systems in place to figure out who the strong performers are and you don’t have mechanisms for teachers to improve, then why would that model work or even survive? Policymakers are always looking for silver bullets, but there don’t appear to be any.”⁹

Fortunately, state and national policymakers are now taking some tentative steps toward encouraging better alignment between teacher compensation reforms and policies related to other human resources areas, including teacher evaluation and professional development. Case in point: Eight states now boast pay-for-performance programs that also incorporate professional development programs for teachers to some extent, as detailed by Robin Chait at the Center for American Progress in her report, “Current State Policies that Reform Teacher Pay.”¹⁰

At the federal level, legislation introduced by then Sen. Barack Obama (D-IL) in 2007 would have required local applicants for “Innovation District” grants to address a range of human resource strategies, including compensation reform to improve the education workforce.¹¹ Similarly, the TEACH Act bills introduced by Sen. Edward Kennedy (D-MA) and Congressman George Miller (D-CA) would require districts to address teacher hiring and placement policies to be eligible for federal grants to offer exemplary teachers “premium pay” to work in high-needs schools.¹² Although the TEACH Act was not passed by Congress, it remains highly influential in conversations about federal policy. For example, in 2007, Miller and Rep. Buck McKeon (R-CA) drew heavily from the TEACH Act in crafting their much-discussed “Discussion Draft” for reauthorizing the Elementary and Secondary Education Act, even lifting parts of it wholesale.¹³

At present, though, the most prominent teacher compensation reform program is the Teacher Incentive Fund, the largest federal program providing targeted support for compensation reform. TIF requires states and local education agencies seeking grants from the fund to address professional development and teacher evaluation when designing performance-based compensation systems. Importantly, TIF is due to expand rapidly this year because of the American Recovery and Reinvestment Act of 2009, which more than doubled the federal appropriation for TIF by allocating \$200 million for the program in fiscal year 2009, which ends at the end of September. What’s more, in May President Obama proposed increasing the program’s appropriation to \$487.3 million in fiscal year 2010, beginning in October.

Even so, the emphasis on the alignment of teacher compensation reform with other human resources reforms is relatively new, and little is understood about what true alignment looks like and how it can best be achieved. One key question: Does TIF encourage actual *alignment* among compensation, evaluation, and professional development, or does it merely require grantees to include “multiple components” without ensuring that various human resources elements truly support and reinforce one another?

This paper reviews emerging ideas about policy alignment in education based on “strategic human resource strategies” in the private sector. Specifically, the paper first examines the Teacher Advancement Program introduced by the Milken Family Foundation in 1999 and now sponsored by the National Institute for Excellence in Teaching—to gain a better picture of what alignment looks like and how it is achieved “on the ground.” We then discuss potential challenges for achieving different kinds of alignment, and then offer recommendations to policymakers interested in encouraging better aligned teacher compensation reforms and other human resources reforms.

The analysis of TAP in this paper suggests that it is possible to tightly align teacher compensation with other human resources reform policies, but that such alignment requires a highly intentional design and cannot be left to chance. The TAP design does not achieve alignment merely by including teacher evaluation and professional development along

with teacher pay in the model, but rather by employing several explicit strategies that allow other schoolwide practices to support and reinforce differentiated compensation, and vice versa. Specifically, the TAP design employs the following methods to ensure an aligned approach to performance-based compensation:

- Teacher evaluation and professional development help teachers develop a clearly defined repertoire of instructional skills that are rewarded by annual bonuses.
- The school's improvement planning process and professional development provide teachers with new instructional strategies that have been proven to produce learning gains for students in the school—another factor rewarded by annual bonuses.
- Differentiated pay is used to create a team of teacher-leaders who have the authority, time, and expertise to *improve* teacher evaluations, professional development, and school improvement planning.

Achieving widespread consensus that traditional ways of paying teachers must change is just the first step on the path to worthwhile reform. Now policymakers are confronting difficult design issues as they craft policies to advance performance-based compensation. So far most of the research and debate has focused on criteria for triggering annual performance bonuses. This paper will illustrate that policymakers must broaden their thinking about compensation reform to consider how other policies can support better ways of paying teachers, and—just as important—how all of these new investments in performance-based compensation can be leveraged to build the capacity of our public schools to take on the hard work of systemic improvement, without which it will be impossible to raise the achievement of America's students to globally competitive levels.

Understanding alignment

With funding from several prominent foundations, the Consortium for Policy Research on Education, or CPRE—a collaborative project of the University of Wisconsin-Madison, the University of Pennsylvania, Harvard University, the University of Michigan, and Stanford University—launched a new project last year called Strategic Management of Human Capital to help state and local policymakers deploy better strategies for recruiting, developing, and retaining talented teachers and principals. One of the project’s primary ambitions is to help policymakers create better alignment among the major personnel-related policies and practices within public school systems.

Based on research and new ideas about human resource strategies in the private sector that have emerged over the last 15 years, CPRE’s Strategic Management of Human Capital project emphasizes two distinct ways to align human resources policies in public school systems. The first is vertical alignment, which refers to the fit between a particular human resources practice, such as employee compensation, and the school system’s overall goals and improvement plan. And the second is horizontal alignment, which refers to the extent to which one human resources policy reinforces and supports another human resources policy. Let’s consider each in more detail.

Vertical alignment

Vertical alignment implies a “good fit” between human resources practices and the larger goals an organization is trying to achieve. A school system can achieve vertical alignment by ensuring that the policies and practices of any given HR area supports the system’s specific student achievement goals and related education improvement plan.¹⁴ For instance, if a district establishes goals for improving student math and reading scores and develops a written plan to reach those goals, then the district’s compensation policies should intentionally be realigned to acquire and develop the teaching talent necessary to achieve those goals.

CPRE’s Heneman and Milanowski suggest that one effective mechanism for securing a good vertical fit is to adopt a written description of the specific teacher competencies necessary to realize the goals of the school or the district and then to ensure that each HR area focuses on maximizing those particular professional competencies.¹⁵ A school or a district could develop its own teacher competency framework or adopt an existing one if

there is sufficient evidence that the instructional skills it describes have a positive impact on student achievement.¹⁶ A number of competency frameworks are available, including Charlotte Danielson's Framework for Teaching and the University of Virginia's Classroom Assessment Scoring System, among others.¹⁷

CPRE's Odden cautions that when considering compensation as one such HR strategy, policymakers need to distinguish between several different elements of teacher pay. Each of these elements should be vertically aligned with the competencies the district hopes to promote so teachers can help students reach the learning goals the district has set. These elements include:

- Base Pay, or the amount teachers can expect to earn in each monthly check
 - Base Pay Progression, or the ways teachers can increase their base pay over time
 - Variable Pay, or any extra amount teachers might earn that is not guaranteed at the beginning of the year, for example through end-of-year performance bonuses.¹⁸
-

Horizontal alignment

Horizontal alignment implies a good fit *among* HR practices. A school system can achieve horizontal alignment by ensuring that its policies and practices in any given HR area support and reinforce the policies and practices in each of its other HR areas.¹⁹ Specifically considering teacher compensation, the question is: Do our policies for paying teachers support and reinforce our policies for recruiting, selecting, inducting, mentoring, evaluating, managing, and providing professional development to our teachers, and vice versa?²⁰

Of course, alignment is not the only HR challenge facing school systems. The quality of basic HR practices is a significant problem, too, as the next two sections on teacher evaluation and professional development will illustrate. In fact, practices in these other HR areas typically are just as outdated and ineffective as current teacher compensation policies.

As this paper will demonstrate, a highly aligned approach to teacher compensation and other human resources policies can actually *create the capacity to improve* multiple HR practices at once. Before examining how that can be accomplished, however, it is important to examine current problems with the two HR functions most often included in initiatives promoting compensation reforms—professional development and teacher evaluation.

The problem with professional development

Most teachers participate in some form of professional development activities every year, at significant public expense. Even a very conservative estimate suggests that the federal government provided well over \$2 billion for professional development in the 2007-8

school year.²¹ Including the cost of collaborative planning time, urban districts can spend as much as \$6,000 to \$8,000 per teacher per year to improve classroom instruction.²² (Although research on professional development spending has focused on urban districts, efficient use of such resources is clearly a concern for suburban and rural districts as well.)

Yet much of what counts as professional development is fragmented and not sufficiently focused on strategies for improving classroom instruction in specific content areas or even linked to school and district improvement plans. Most states mandate that teachers accumulate “professional development credits” every five years to renew their licenses—a requirement teachers often fulfill by taking university courses of questionable relevance or quality that have little or no impact on teaching effectiveness.²³

Moreover, many teachers still attend short-duration “one-shot” workshops that offer no follow-up to help them implement new strategies in the classroom.²⁴ Last year, federal funds allocated under Title II, Part A of the Elementary and Secondary Education Act paid for more than 1.9 million teachers to attend one-day workshops that took place outside of the school day.²⁵

Nearly every district also provides teachers with workshops during school hours as part of contractual “in-service” days. But researchers have observed that teacher absenteeism increases on those days because teachers themselves do not find the workshops very relevant or useful.²⁶ Indeed, according to a federal survey of teachers conducted in 2003-4, only 59 percent of teachers who participated in professional development focused on the content areas they teach found such training “useful” or “very useful,” while fewer than half considered the training they received on other topics to be useful.²⁷

So far, few rigorous research studies find strong links between professional development and student achievement. A 2007 review of more than 1,300 studies on professional development identified only nine that met standards for scientifically based research established by the federal What Works Clearinghouse, a project of the U.S. Department of Education that reviews research evidence on specific interventions to improve student outcomes. Among the nine studies, no professional development training lasting 14 or fewer hours had a positive impact on student achievement; in contrast, professional development of extended duration (an average of 49 hours) boosted student achievement by about 21 percentile points.

Unfortunately, the effective programs varied too little in design to determine if any specific aspects other than duration were particularly effective.²⁸

Despite the paucity of rigorously scientific studies, education experts in this policy arena argue that professional development should include the following elements:

- Focus on actual course content
- Provide opportunities for active learning rather than just passive listening

- Require the collective participation of teachers from the same grade, department, or school
- Seem coherent to teachers because it directly relates to their own day-to-day classroom work, their schools' particular improvement plans, and their states' academic standards and assessments
- Be of sufficient duration—sustained over time and involving a substantial number of hours—to actually influence teacher practices.²⁹

Two influential studies conducted in conjunction with the Eisenhower Professional Development Program—the federal government's major investment in teacher professional development under the Elementary and Secondary Education Act prior to its reauthorization by the No Child Left Behind Act of 2002—found evidence that these elements had a positive impact on teachers' knowledge and classroom practices.³⁰ Yet the first of those studies, which was based on a national sample of teachers in 93 percent of U.S. districts, found that few teachers experienced high-quality professional development incorporating those elements. The average activity lasted less than a week and the median participant spent only 15 hours engaged in it. Most activities did not involve collective participation by teachers who worked together, did not focus significantly on content, exhibited little coherence, and offered few opportunities for active learning.³¹ Many experts also recommend that professional development should:

- Be based on and include ongoing analysis of student performance data
- Be part of teachers' everyday work rather than an "add on" to it
- Provide follow-up and coaching to help teachers implement new practices in the classroom
- Include some form of accountability to ensure that new practices are implemented appropriately
- Include methods to evaluate whether the professional development had a positive impact on student learning.³²

Since all of those recommendations can best be realized by locating the site of training and support closer to teachers' daily work, experts call for professional development that is "intensive, sustained, and job-embedded."³³ In this way, teachers would have common planning time and the ability to work with each other and with experts in their own schools and classrooms.

National data suggest that many teachers do participate in at least some activities that offer ways to collaborate on the job with each other and with experts to improve instruction. According to the federal government's *2003-4 Schools and Staffing Survey*, 70 percent of U.S. teachers reported participating in regularly scheduled collaboration with other teachers on issues of instruction during the previous 12 months; 63 percent reported observing or being observed by other teachers; and 46 percent received or provided mentoring or coaching.³⁴

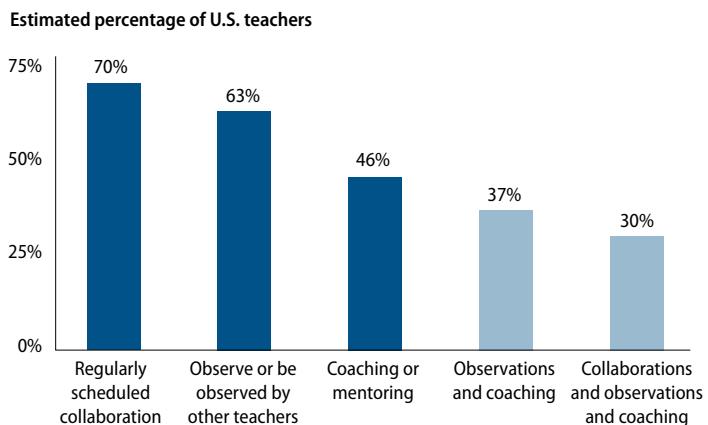
In the same SASS survey, nearly a third of teachers who participated in collaborative planning time reported no classroom observations whatsoever.³⁵ Of course, even when teachers have opportunities for observations as well as common planning time, they might not know how to observe classroom teaching with enough objectivity and specificity, or they might feel uncomfortable bringing up such evidence when they meet with each other.³⁶ Ineffective collaboration time bears huge fiscal as well as opportunity costs. According to studies by Education Resource Strategies, a Massachusetts-based organization that helps school districts around the country use resources more strategically, collaborative planning time can account for approximately 65 percent of total spending on professional development at the school building level.³⁷

Yet it is not clear these activities are conducted in useful ways or complement each other enough to form a coherent professional development program. Case in point: Only an estimated 30 percent of teachers participated in all three kinds of activities during the year in question (see Figure 1).

In their new book *Instructional Rounds in Education*, Elizabeth City, Richard Elmore, and Lee Tielert of Harvard University, and Sarah Fiarman, principal of the Martin Luther King, Jr. School in Cambridge, Massachusetts note that teachers who only have common planning time but do not critically observe each others' instruction might not be able to link problems with student achievement back to problems with classroom instruction, resulting in "misiagnoses" that waste professional development time. In their example below, teachers identified a need to seek out remedial strategies as the next step for professional development rather than attending to more basic issues of classroom instruction that, if addressed as the next step, might obviate the need for remedial strategies:

We then observed a team meeting of teachers in the grade level whose classrooms we had observed. ... In the meeting a problem emerged. The student work was obviously quite variable from classroom to classroom. ... Each teacher offered an explanation [that] mainly had to do with the teachers' interpretations of the students' skill levels at the beginning of the unit. ... So the discussion quickly shifted to what kinds of remedial strategies one might use ... What the teachers didn't know—because they had never observed each other teaching—was that ... the variability in student performance was a result of the teaching that was going on and the actual tasks that students were asked to do, not, as the teachers hypothesized, a result of students' prior knowledge.³⁸

FIGURE 1
Teachers who experience various elements of "job-embedded" professional development



Source: U.S. Department of Education 2003-4 Schools and Staffing Survey.

Note: First three bars based on data from NCES Data Analysis System (DAS), <http://nces.ed.gov/das/>. Fourth and fifth bars reflect author's personal estimates based on data from DAS.

The example illustrates that providing teachers with common instructional planning time might not be an effective way to improve teaching unless it is coupled with strategies that focus on assessing and improving the quality of instruction in teachers' individual classrooms.

In an effort to address problems with traditional professional development and respond to federal accountability and school improvement requirements, schools and districts are increasingly turning to teacher mentors and instructional coaches.³⁹ A number of large urban districts, including New York, San Diego, Boston, Dallas, Philadelphia, and Chicago, have made coaching a central part of instructional improvement efforts.⁴⁰ The trend is not limited to urban school systems. For example, officials at the North Dakota Department of Education have noted a "tremendous interest" in using instructional coaches and have issued guidance on using federal funds to pay for them.⁴¹

The idea is to provide teachers with a form of professional development that can be individualized to their particular needs and to bring expertise directly into the classroom to help teachers implement new practices.⁴² Yet the research on implementation and success of coaching remains thin.⁴³ And several new studies raise serious concerns about the limitations of coaching as it is currently practiced in schools.

An in-depth, three-year study of instructional coaching in a sample of Urban Systemic Initiative sites with grants from the National Science Foundation found that coaches focused almost exclusively on "show and tell" to describe and model effective strategies, offering little to no "hard feedback" to classroom teachers that could help them reflect on and change their own instructional practices. Instead, coaches provided only soft feedback, non-controversial comments carefully crafted not to offend, "offering help and encouragement, but ignoring bad practice."⁴⁴

As a result, coaching seldom moved beyond the relationship-building phase and had limited impact on instruction. "The crux of the one-on-one work appears to lie in the structure of teacher-leaders' feedback to classroom teachers," the authors of the NSF study concluded, "and here the overreliance on soft feedback can be crippling when bold changes are envisioned."⁴⁵

An experimental study on the effects of a rural, NSF-funded program that included "peer coaching" identified a similar problem. "Peer partners provided minimal critique to the classroom teachers whom they observed," the researchers found. "Overall, peer partners did not challenge or question one another's classroom practices." Instead, "discussions between peer partners consisted of explanations of what occurred during the classroom observations rather than meaningful analyses of how classroom instructions could be improved." Not surprisingly, the study found that peer coaching resulted in no improvement in students' mathematics achievement.⁴⁶

The importance of time and feedback

The disappointing findings about misspent professional development time and the lack of hard feedback in coaching models might go a long way toward helping us understand why even professional development that *appears* to be “intensive, sustained, and ongoing” might not actually improve instruction. Evidence from outside of education suggests that well-spent time and objective feedback are critical for improving professional performance, particularly for those who have more experience.

Cognitive scientists, for example, find that for most individuals performance improves rapidly during their novice period but then levels off dramatically, often for the rest of their careers—a finding that matches what education researchers have found in “value-added” studies of teacher effectiveness.⁴⁷ After individuals reach an adequate or acceptable level of performance, additional experience does not translate into greater effectiveness.

The reason is that the mental procedures for performing tasks become “automated,” and the instructions for them are relegated to long-term memory, which allows the conscious mind (what cognitive scientists call “working memory”) to focus on other problems.⁴⁸ The same is true of everyday tasks, which is why people can think about other things while driving a car, and also why their driving skills do not improve despite decades of additional experience behind the wheel.⁴⁹

In every career field, however, some individuals do continue to improve and reach high levels of expert performance. By studying elite performers whose trajectories do not match the typical “arrested development” pattern, cognitive scientists have identified one key to continuous improvement—something they call “deliberate practice.” Deliberate practice is different from the common notion of practice as short bursts of mindless, repetitive activity. It requires a great deal of time and intense, often exhausting, mental concentration.⁵⁰ According to this new understanding of expert performance, people only improve at any particular skill when they:

- Make a conscious effort to focus “working memory” on *improving* performance
- Obtain critical feedback from an objective expert in the field who can identify strengths and weaknesses and specify concrete skills to target for improvement
- Engage in activities designed to improve those skills while reinforcing areas of strength.

Moreover, individuals only reach *high* levels of expertise by repeating that cycle for a number of years, continually building their repertoire of strengths by identifying more complex skills to master.⁵¹

Importantly, when cognitive scientists talk about the “automaticity” that sets in for many individuals, they are not implying that professionals do not work hard at their jobs. Indeed,

most teachers work very hard at their jobs. That is the problem: To significantly improve one's level of performance, professionals must find *additional* time, outside of day-to-day practice, to focus and concentrate on improvement *per se*. Time spent doing a task is not necessarily time spent improving at the task. A professional can work exhausting, 80-hour weeks for years while never improving his or her skills in the job.

That creates a problem for American teachers. One recent study found that U.S. teachers spend about 80 percent of their total working time engaged in classroom instruction, during which they have to focus mental energies on doing the job rather than improving their ability to do the job, compared to about 60 percent for teachers in a number of high-performing countries.⁵² And much of the 20 percent of time that is left must be spent on any number of tasks such as grading papers, analyzing student data, planning lessons, and completing paperwork.

On top of that, just at the point when teachers exit their novice periods and begin to "automate" basic classroom management and instructional tasks, many teachers face additional demands which compete for time that could be invested in deliberate practice. A recent study of such "second-stage" teachers by Cheryl Kirkpatrick, a researcher with Harvard University's Project on the Next Generation of Teachers, found that while most wanted to improve their instruction, they faced requirements to obtain master's degrees and renew their teaching licenses by accumulating professional development credits. "These demands, ironically, actually seemed to draw participants away from investing actively in their teaching," Kirkpatrick concluded.⁵³

K. Anders Ericsson, co-editor of *The Cambridge Handbook of Expertise and Expert Performance* and a leading authority in the field, observes that most workplaces are not designed to encourage or support improvement. "The lack of paid time to practice raises interesting challenges for professionals who want to improve their performance," he writes in a new book, *Development of Professional Expertise*, to be published later this summer.

Lack of time, however, is not the only challenge. According to Ericsson, "The greatest obstacle for deliberate practice during work is the lack of immediate objective feedback."⁵⁴ Lack of feedback also arose in the study of second-stage teachers. According to Kirkpatrick:

When second-stage teachers chose to invest in their teaching, they did so in ways that were not necessarily informed by careful analysis of their practice and how they could improve it. Many wanted to teach better. However, their [schools] did nothing to help them realize what they had to learn or what they might stand to gain from different types of investment.⁵⁵

Adding lack of time and feedback to the problems discussed earlier, it becomes clear that current approaches to professional development suffer from a great many limitations:

- Too much professional development consists of “one-shot” workshops and other activities of brief duration.
- Too many activities focus on general topics such as classroom management rather than specific content in the curriculum for which teachers are responsible.
- Too many activities are not based on analysis of students in the teachers’ school or classroom so they do not focus on students’ specific learning needs.
- Too many activities take place after school and away from school, removed from where the work happens.
- Too many activities focus on “generic” topics or strategies that do not relate to teachers’ day-to-day work and specific challenges.
- Too many activities consist of “sit and get” passive listening rather than active learning to solve problems.
- Teachers attend too many activities alone rather than with colleagues in their own departments, grade levels, or schools.
- Teachers have little time to work on improving their professional practice, face competing demands in the form of meaningless certification requirements, and often do not have the structure and support to use existing collaborative time well.
- Too little professional development offers follow-up opportunities to help teachers figure out how to implement new strategies in their own classrooms (and to make sure they do).
- Even teachers who have instructional coaches rarely receive the kind of honest, specific “hard feedback” on their own classroom teaching that could help them identify strengths and weaknesses so they can improve day-to-day instruction.

Of course, when it comes to the last of those problems, the unwillingness or inability of peers and instructional coaches to provide hard feedback might not matter so much if the formal teacher evaluation process filled the gap. But it turns out that teachers are even *less likely* to receive constructive, critical feedback as part of their periodic job reviews.

The problems with teacher evaluation

Several recent reports have drawn attention to the low quality of teacher evaluations in the United States, calling the process cursory and capricious and the results meaningless. “The troubled state of teacher evaluation is a glaring and largely neglected problem in public education, one with consequences that extend far beyond the performance pay debate,” concluded Thomas Toch, former co-director of Education Sector, and Robert Rothman, a principal associate for the Annenberg Institute for School Reform at Brown University, in a study released by Education Sector last year.⁵⁶ The problems are many:

- **Evaluations are infrequent.** According to collective bargaining agreements, fewer than half of the nation’s 50 largest school districts require tenured teachers to be evaluated even once per year, and only about a quarter of the districts require untenured teachers to be observed two to three times per year⁵⁷ (see Figure 2). Only 29 states require evaluations to be based on *any* classroom observations.⁵⁸

- **Evaluations are based on scant evidence.** In a typical district that requires only one observation per year lasting a minimum of 45 minutes, a tenured teacher's evaluation would be based on information collected during less than one-tenth of one percent (0.08 percent) of her teaching time.⁵⁹ In addition, only 15 states require teacher evaluations to consider objective measures of student learning.⁶⁰
- **Evaluations rely on crude instruments.** Evaluation criteria often take the form of crude "checklists" focusing on teacher behaviors and attributes that are not necessarily correlated with student learning. According to Michigan State University's Mary Kennedy: "It's typically a couple of dozen items on a list: 'Is presentably dressed,' 'Starts on time,' 'Room is safe,' 'The lesson occupies students.'"⁶¹ Many instruments are equally crude in how they report results: A study of teacher evaluation policies in 12 districts across four states—Arkansas, Colorado, Illinois, and Ohio—by the New Teacher Project, an organization that helps school districts improve teacher recruitment and placement, found that five in ten districts allowed principals only to rate teachers in a binary fashion as either "satisfactory" or "unsatisfactory."⁶²
- **Many principals are inadequately trained.** Across 12 districts studied by the New Teacher Project, only half of principals surveyed said they had received "extensive" or "very extensive" training on how to conduct an effective evaluation.⁶³ Using only one evaluator who has not been very well trained also contributes to charges of subjectivity and lack of reliability in evaluations.
- **There are no reliability checks.** Typically, principals alone are responsible for evaluating teachers, without "reliability checks" to ensure that their judgments are fair and objective.⁶⁴
- **The results are inflated.** The New Teacher Project study found that in five districts using binary ratings, more than 99 percent of teachers were judged "satisfactory." In five districts where a broader range of ratings are available to principals, more than 70 percent of teachers received the highest rating of "outstanding," "superior," "distinguished," or "excellent," while 24 percent received the second-highest rating. In Denver and Cincinnati, less than 10 percent of low-performing schools rated even one tenured teacher "unsatisfactory." Those ratings contrast sharply with findings from a recent study of over 800 first-grade classrooms nationwide, which rated only 23 percent as exhibiting "high overall quality" based on teachers' instructional practices and ability to establish a supportive climate.⁶⁵
- **Expectations are inflated.** The same study also found that, in six districts where teachers could be rated at multiple levels, half of both the tenured and untenured teachers who did *not* receive the very highest rating thought they deserved to do so. The authors concluded such attitudes are understandable in the current system, which creates a reinforcing loop where inflated ratings breed inflated expectations, which in turn cause principals to feel pressured to inflate ratings.⁶⁶

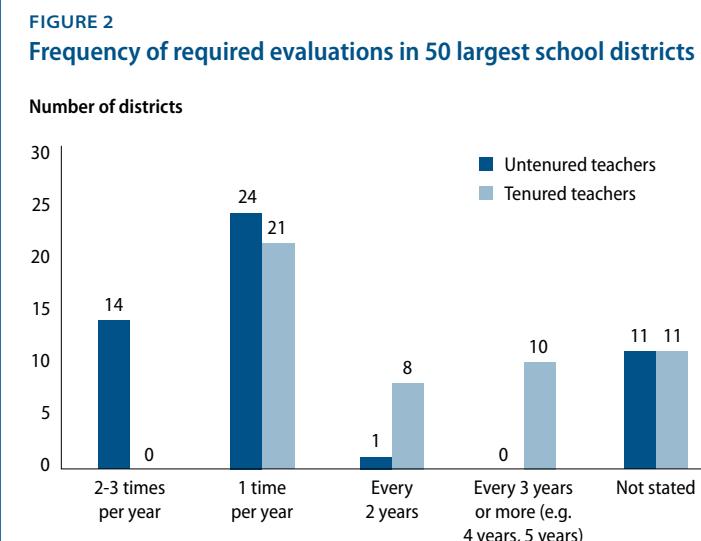
- **The results are not used for performance management, creating tremendous HR misalignment.** Perhaps due to their poor quality, teacher evaluations are not used for performance management purposes, and they are disconnected from nearly every other human resource function in school systems. Across the 12 districts studied by the New Teacher Project, no district used evaluation results to inform recruitment, professional development, retention, or layoffs. Only one district linked evaluation to hiring and placement, one linked it to tenure, and one linked it to compensation.⁶⁷
- **Teachers receive little or no useful feedback to improve their performance.** Across the 12 districts studied by the New Teacher Project, evaluations identified areas for improvement for only 26 percent of all teachers and only 43 percent of novice teachers. “It is inconceivable that 74 percent of teachers, and 57 percent of teachers in their first three years, do not require improvement in any area of performance,” the authors of the study remarked.⁶⁸ It should come as no surprise that, in a recent national survey, 41 percent of teachers said evaluation “was just a formality” and 32 percent said it was “well-intentioned but not particularly helpful” for improving their teaching practice.⁶⁹

Of course, principals might provide feedback informally outside of the evaluation process, but for many teachers that does not seem to be the case. In the New Teacher Project study, 47 percent of teachers across the 12 districts said they did not have even one informal conversation with an administrator about improving aspects of their instructional performance over the past year.⁷⁰ That tracks closely with the federal government’s 1999-2000 survey in which only 11 percent of U.S. teachers strongly agreed that “the principal talks with me frequently about my educational practices.”⁷¹

Taken together with the problems in professional development discussed above, such findings paint a grim picture of the education system’s capacity to encourage and equip teachers to improve their practice on any regular basis. Teachers are misled into believing they do not need to invest the time and effort to improve. They receive little encouragement to do so. And they are not rewarded for doing so.

Perhaps worst of all, when teachers decide to make the investment anyway, professional development and evaluation systems seem designed to withhold rather than provide them with the one resource that cognitive scientists say is essential for improving performance—objective feedback.

In April, U.S. Secretary of Education Arne Duncan signaled an intention to draw attention to the prob-



Source: Jane Hannaway and Andrew J. Rotherham, “Collective Bargaining in Education and Pay for Performance” (Nashville, TN: National Center on Performance Incentives, 2008). Based on analysis of collective bargaining agreements.

lem of teacher evaluation in a letter to governors about potential metrics that the department might collect under the American Recovery and Reinvestment Act.⁷² Later that month, the department issued a list of suggestions for using stimulus funds to improve education, including one related to teacher evaluation:

Establish and implement a fair and reliable teacher evaluation system that provides ongoing feedback to teachers about their performance based on objective measures of student achievement outcomes and multiple classroom observations, that gives guidance for improving instructional practices, and that is used to inform teacher professional development and advancement.⁷³

But an important question remains: Even if states and districts wisely invest ARRA funds to develop better evaluation systems, does the education system itself have the capacity to implement such systems on an ongoing basis?

Confronting the capacity challenge

Assume that an elementary school with 25 classroom teachers wants to implement a rubric-driven, multiple-observation approach to teacher evaluation. If the principal is to conduct three observations per veteran teacher, and five observations for the four novice teachers the school has just hired, that will translate into 83 classroom observations and 83 post-conferences over the course of the school year, which typically includes only 180 instructional days to begin with.

On top of that, the principal would need to find time to complete paperwork required by the new system. And that does not include pre-conferences before announced observations, as some standards-based evaluation models require. Not surprisingly, studies of local attempts to implement such standards-based evaluation systems have raised serious questions about organizational capacity:

- A study of a medium-sized Midwestern district piloting an evaluation system based on Danielson's Framework for Teaching, with most teachers receiving one observation and a subset receiving six, found that administrators had trouble handling the workload. Some failed to complete all of the observations, and some even failed to produce final ratings. In addition, many administrators told researchers they did not have the time to provide extensive feedback or coaching to teachers based on their observations.⁷⁴
- A study of three districts in different parts of the country implementing new evaluation systems based on the Framework found that even though two districts continued to stagger evaluations of tenured teachers over three- and four-year periods, "principals and other evaluators in some schools clearly struggled with managing the new system on top of ... already substantial burdens administrators faced in leading school

organizations.” As a result, most principals had to sacrifice personal time to complete evaluations, and some administrators cut corners. Feedback provided to teachers was largely affirmative rather than deep or critical, and veteran teachers did not deem it useful for changing practice.⁷⁵

- A study of a large Western school district implementing a new evaluation system based on the Framework, with one observation for tenured teachers and nine for novices, found that the system absorbed as much as 25 percent of principals’ time and forced them to sacrifice large amounts of personal time to complete evaluations. Principals also cut corners by conducting only brief classroom visits and engaging teachers in very quick chats to satisfy post-conference requirements. Moreover, researchers concluded that principals lacked the skills or the will to provide specific, critical, and valuable feedback to veterans to help them improve their performance.⁷⁶

Indeed, the last study suggests that new standards-based evaluation systems can simply reproduce many of the same tepid outcomes as traditional evaluations: “Principals emphasized praise in written evaluations and provided ‘gentle’ criticisms if they criticized teachers at all. …Very little critical feedback was provided either through evaluation scores or in narratives.”

“Principals did not assign an unsatisfactory rating in any of the 485 written evaluations we reviewed,” despite the fact that in interviews principals did refer to instances of sub-par teacher performance.⁷⁷ Clearly, simply adopting a standards-based evaluation model is no guarantee that evaluations will improve significantly in either rigor or usefulness, despite significant investments of time on the part of administrators.

Yet in a study for Education Resource Strategies published last year, ERS director Regis Anne Shields and executive director Karen Hawley Miles found that it is possible for schools to implement a much more rigorous and dynamic evaluation teacher evaluation system that can inform both professional development and performance management. Shields and Miles examined nine small urban high schools dubbed “leading edge” because they attain better performance outcomes by organizing time and resources differently. The analysis concluded that such schools could only implement better evaluations because they had achieved a “limited span of review,” in other words, a much smaller than average ratio of teachers per evaluating administrator.⁷⁸

“In a typical large high school of 1,500 students, a principal and assistant principal(s) may be responsible for evaluating anywhere from 35 to 130 teachers each, depending on the number of assistant principals and their spheres of responsibility,” Shields and Miles observe. However, “the Leading Edge Schools have small spans of review, ranging from four to 34 teachers. As a result, leaders at these schools are able to evaluate teachers at least once a year and with much more focus and knowledge of the teachers’ practice and skills.”

Importantly, the schools' manageable spans of review are not just a positive side effect of their small enrollments, the authors discovered. Rather, they are made possible through budgetary tradeoffs that allow the schools to spend a much higher proportion of their per-pupil dollars on administrative leadership.⁷⁹

Some reformers have suggested that other schools can build capacity for better evaluations by hiring a co-principal so that one can provide instructional leadership while the other manages school operations.⁸⁰ But it is important to be clear that simply adding another administrator might not automatically allow schools to implement an evaluation process based on "multiple classroom observations" per year. Despite strategic investments in leadership capacity, for example, only in three of the nine leading-edge high schools studied by Shields and Miles do administrators observe every teacher more than once per year.

Toch and Rothman offer another suggestion: "Because principals lack the time and the training to conduct comprehensive teacher evaluations ... school systems should create cadres of trained district-level evaluators of the sort that Toledo has established under its peer review program."⁸¹ Toledo's program, instituted in 1981, uses experienced teachers called Intern Consultants to mentor and evaluate teachers during their first year working in the school district. At the end of the year, Intern Consultants make recommendations for continuing employment or dismissal to an Intern Board of Review composed of four administrators and five teachers.⁸²

However, such peer assistance and review programs only provide rigorous evaluations for a small subset of a district's teachers, usually only novices plus a handful of experienced teachers with serious problems. In Toledo, the latter group typically amounts to no more than two to three veterans per year.⁸³ If the aim is to conduct better evaluations of all teachers on a more regular basis, then it is not clear whether and how the PAR approach can be expanded to achieve that goal.

Tapping the power of alignment

How an integrated approach to compensation reform can solve multiple problems

The Teacher Advancement Program aligns compensation vertically with school improvement planning and horizontally with professional development and evaluation.

Therefore, TAP offers an important opportunity to study what alignment might look like and how it can be achieved in a design that actually has been implemented “on the ground.” Of course, TAP is not the only possible way to design compensation reforms to align with other HR components. But TAP is one of very few examples of an “aligned design” that actually has been implemented in a large number of schools.⁸⁴

Last year TAP, which was introduced by the Milken Family Foundation in 1999 and now is sponsored by the National Institute for Excellence in Teaching, operated in 219 schools serving over 70,000 students and employing over 6,000 teachers in 14 states and the District of Columbia.⁸⁵ Two states, South Carolina and Louisiana, have adopted the TAP design for voluntary performance-pay initiatives in high-needs schools and provide technical support for the program at the state level.

Moreover, in 2005 Minnesota passed legislation establishing a statewide “Q-Comp” program offering funds to districts to design compensation reform programs based on guidelines strongly inspired by the TAP design. Last year 39 Minnesota school districts and 21 charter schools participated.

TAP in Action

The National Institute for Excellence in Teaching, or NIET, emphasizes four essential components for any school adopting the program:

- Multiple career paths
- Instructionally focused accountability
- Performance-based compensation
- Ongoing, applied professional growth

In this section of the paper, we will examine each of these components in turn.

Multiple career paths

TAP allows teachers to take on additional professional roles and responsibilities as they progress in their teaching careers. “Career Teachers” are full-time classroom teachers. “Mentor Teachers” teach students part-time but also provide instructional coaching to a group of Career Teachers under their supervision, observing and providing feedback, modeling strategies, team teaching, planning lessons, and working with teachers on their Individual Growth Plans, documents that list each teacher’s goals for improving instruction and student performance and describe progress toward meeting those goals. Mentor Teachers also conduct formal classroom observations as part of the teacher evaluation process.

“Master Teachers” may continue to teach one or two classes but also oversee the work of Mentor Teachers and share significant authority with the principal for schoolwide improvement, such as analyzing student achievement results to specify learning goals, developing the school’s improvement plan, and identifying, field testing, and overseeing the implementation of new instructional strategies. Master Teachers also may lead teachers’ collaborative planning time, monitor both student and teacher growth, and manage the teacher evaluation process to ensure both reliability and validity.

Like Mentor Teachers, Master Teachers also can provide one-on-one support to Career Teachers through observation and coaching, modeling, demonstration lessons, or team teaching. Mentor Teachers and Master Teachers both serve on a formal School Leadership Team along with the principal and assistant principal.

Instructionally focused accountability

TAP uses this term mainly to refer to its human resources practices related to evaluation and performance management. Each teacher is observed four to six times per year during announced and unannounced visits by administrators, Master Teachers, and Mentor Teachers, all of whom have been trained and certified to evaluate teachers on a scaled rubric based on TAP’s Teaching Skills, Knowledge and Responsibilities Performance Standards. “The rubric,” as it is more commonly known in the program, incorporates some elements of Charlotte Danielson’s Framework for Teaching, but the program has expanded Danielson’s Framework and added more detail to it, along with stretching the scoring system one more notch so that teachers can score from 1 to 5 rather than only up to 4.

Thus, teacher evaluation in TAP is part of a comprehensive system of “performance management.” Prior to announced observations, the evaluator meets with the teacher in a pre-conference, and all observations are followed by a post-conference. During the post-conference, the teacher receives feedback related to all areas of the rubric, including specific areas for “reinforcement” and for “refinement,” as well as concrete advice on next steps to improve his or her practice. The results inform an Individual Growth Plan for each

teacher. Also, through the annual bonuses described below, teachers are compensated in part based on the scores they receive on the evaluations.

Evaluation results also directly inform professional development. Master and Mentor Teachers follow up on evaluations to provide targeted coaching and support to teachers as they work on areas identified for refinement. Also, because all evaluation results are entered into a data system called CODE, Master Teachers can analyze which areas of the rubric have the lowest scores faculty-wide and address those skills in weekly cluster meetings.⁸⁶ TAP’s version of teacher evaluation thus directly addresses the problems with traditional evaluation described earlier:

- **Problem: Evaluations are infrequent, based on scant evidence, and rely on crude instruments.** In TAP, every teacher is evaluated every year, and teachers’ evaluations must be based on at least 4 to 6 observations per year, some announced and some unannounced. The TAP rubric is a far more extensive and detailed description of good teaching than the “checklists” typically used for teacher evaluations. A study published in the *Economics of Education Review* found teachers’ ratings on the TAP rubric to be associated with higher student achievement gains.⁸⁷
- **Problem: Many principals lack expertise, and there are no reliability checks.** Principals, Master Teachers, and Mentor Teachers all receive rigorous training to conduct evaluations using the TAP rubric, and they must be recertified on a regular basis. Teachers are observed by multiple trained evaluators every year, and the School Leadership Team is specifically tasked with ensuring high levels of inter-rater reliability.
- **Problem: The results are not used for performance management, and teachers receive little useful feedback to improve their performance.** In an aligned model such as TAP, evaluation results are used to inform teacher compensation and to plan individual and schoolwide professional development. Teachers receive extensive and highly detailed feedback on their performance, including information on specific areas for “reinforcement” and for “refinement,” along with advice on how to improve.
- **Problem: The results are inflated, and most teachers expect the highest rating.** TAP’s rubric is scaled from 1 to 5, with “3” formally designating satisfactory performance and “5” designating rare, exceptional performance. While NIET does not publish aggregate evaluation results, anecdotal evidence suggests the program takes pains to enforce rigor. “Three means that you’re proficient,” explains Monica Knauer, a Master Teacher at Dwight D. Eisenhower Academy of Global Studies in New Orleans. “It means that you’re a good, strong, solid teacher, and any parent who has a child in your class should be happy.” In contrast, “a five is described as walking on water,” she adds. “If you scored a five in ‘presenting instructional content,’ that means anyone who wants to know how to do that better should come observe you. You are the poster child.”⁸⁸

These standards challenge the inflated expectations created by the current system, according to Eric Matheson, principal of a South Carolina elementary school. “When we started out, teachers immediately thought, ‘How do I get a five on everything?’ But the state TAP folks said, ‘That’s like walking on water.’ It was hard for [teachers] to understand given the way evaluations had worked in the past. I had a lot of conversations where I said, ‘That would mean there’s no room for improvement. Can you honestly say that is true every day for every student on every objective, and that it would not be possible to improve in any area?’”⁸⁹

Moreover, school leadership teams use multiple data sources to check for validity and ensure that evaluation ratings do not become inflated. “When the rubric is applied accurately to teacher instruction, there should be a strong correlation between value-added scores and teacher evaluation scores. We know that from early research on the TAP rubric, so schools look for that,” explains Melodie Barron, a Senior Program Specialist at NIET.

Her example: “This teacher was scoring fours and fives on the rubric, but in the value-added data just a two. They go back and really analyze what’s happening there. If she got high value-added scores in reading, but a lower value-added score in math, was the rubric applied to evaluate teaching in both subjects equally? The two sources of data act as a quality control and help us make sure we are applying the rubric consistently and evaluating teachers validly and reliably.”⁹⁰

The TAP design suggests that a robust evaluation system should benefit classroom teachers, not just supervisors. Indeed, TAP Master Teachers describe an approach that tracks closely with the optimal coaching scenario described by cognitive scientists who study expert performance: observe, provide specific feedback, identify areas of strength to reinforce and a skill for improvement, and offer activities specifically designed to improve that skill.⁹¹ “It’s not just, ‘Oh, you didn’t fare so well in the area of questioning,’ but rather, ‘these are some examples of how you can increase your performance within this indicator,’” explains Alma Velez, a Master Teacher at Jones Elementary School in Bryan, Texas. “I will provide them with activities they could have implemented within that lesson itself to improve in the area of questioning. And then I ask them, ‘in future lessons, how do you see these activities playing out within your lessons?’”⁹²

TAP’s evaluation system accomplishes something else cognitive scientists say is important for continuous improvement—setting a benchmark for expert performance high enough that practitioners across the full continuum of expertise can get feedback and continue to grow. “Prior to TAP, I can honestly say that in the 25 years in this school, there were some years where no one in an administrative or certainly in an instructional leadership role would ever walk in to evaluate me,” says Lynn Kuykendall, a Master Teacher at Clinton Elementary School in South Carolina. “When I would question that, they’d say, ‘Why? You’re doing a great job. You’re nationally board certified.’ But that doesn’t mean I shouldn’t be evaluated. How do you know? And there were personal goals I wanted to work on so that I could keep improving. But if I went a whole year without feedback from an evaluation, it would be hard to work on those goals.”⁹³

Performance-based compensation

TAP's design includes two kinds of differentiated pay for teachers. First, TAP provides annual salary "augmentations" for Master Teachers and Mentor Teachers, who assume leadership roles and additional responsibilities beyond their own classrooms and work a longer school year than the typical classroom teacher. Second, TAP offers all teachers, including Master and Mentor Teachers, a variable annual bonus based on three weighted factors:

- The scores they earn on the performance evaluation described above
- The learning gains of students in their own classrooms based on a value-added model
- Schoolwide learning gains also based on a value-added model.⁹⁴

Thus, TAP improves compensation by paying teachers for performance, for knowledge and skills, and for additional roles and responsibilities. NIET points out that in states that implement TAP in high-poverty schools, the program also addresses the fourth kind of compensation reform—recruitment incentives for "hard-to-staff" schools.

While the additional "roles and responsibilities" compensation for Master Teachers and Mentor Teachers is actually an annual augmentation rather than a reform to the salary schedule *per se*, it means TAP addresses what the Consortium for Policy Research in Education's Odden would call "base pay" and "base pay progression" in addition to bonuses or "variable pay." Thus, TAP illustrates what might be possible if salary schedules rewarded teachers for developing expertise and contributing that expertise for broader improvement efforts rather than only for accumulating experience and education credits. This compensation component turns out to have major implications for alignment—and improvement—of other HR practices.

Ongoing, applied professional growth

TAP uses this term to refer to its version of job-embedded professional development. Led by Master Teachers and Mentor Teachers, professional development in a TAP school includes a collaborative component in the form of weekly "cluster group" meetings, where teachers who work in the same grade levels or subject areas can work on teaching skills related to the rubric, learn new instructional strategies, analyze student work and achievement data, and plan for instruction.

The program also includes an individualized component as Master and Mentor Teachers work one-on-one with Career Teachers to provide them with ongoing coaching and support in their own classrooms—for example by providing a demonstration lesson using a new instructional strategy introduced in that week's cluster meeting or by modeling a skill identified for "refinement" after the teacher's last formal observation. Importantly, the new instructional strategies introduced during cluster meetings are not just "best practices"

brought back from a conference, but rather carefully identified and adapted strategies that relate directly to the school’s improvement plan. This lesser-known element of TAP is critical to understanding how the design achieves very tight vertical alignment.

The process works like this: Master Teachers lead an effort to analyze student achievement data from the previous year in order to identify and target areas of weakness, such as reading comprehension. Based on that analysis, they work with other members of the School Leadership Team to develop an “academic achievement plan” for the coming year. Next, Master Teachers conduct research to identify potential strategies for improving student achievement in the targeted areas—reviewing the relevant research, consulting experts, accessing expertise within the TAP network.

Once Master Teachers have identified a promising strategy, they adapt it as necessary to ensure that it fits the state standards and school curriculum, and then “field test” it with students in the school, perhaps in a Mentor Teacher’s classroom or with students of a Career Teacher who can observe so the Master Teacher can also model a particular skill on the rubric.⁹⁵

When assessing the success of a field test, Master Teachers take into account two factors. The first is the experience of implementing it in the classroom (does this “work with” our students in this school?). The second is student achievement data from formative pre- and post-assessments administered as part of the field testing (does this “work for” our students in this school, by raising their skills?). Moreover, Master Teachers pay attention to how well it works for different groups of students, including low, average, and high achievers. Master Teachers conduct as many field tests as necessary, in as many classrooms as necessary, to fine-tune the strategy and verify its effectiveness for different groups of students in the school.

Field testing serves multiple purposes. First, it allows strategies to be adapted to “fit” the needs of the school’s own students and teachers rather than “just being dropped into classrooms,” a common complaint among teachers in many schools. Second, it offers an assurance that the strategy can, if implemented well, significantly raise student achievement, thus providing teachers with strategies that should boost classroom and schoolwide value-added gains to which annual bonuses are tied. Third, it contributes greatly to the credibility of Master Teachers and Mentor Teachers when they are ready to introduce the strategy to Career Teachers during a weekly cluster meeting.

“It validates everything you provide the teachers,” says Cassandra Slayton, a Master Teacher at Terra Vista in Lubbock, Texas. “If you don’t know how to implement it, and you don’t have the data to support it, why should they buy into it? But since we did test it with students and we did see the gains, we give them all of that data. And we tell them what we saw happen with real students. It’s not a made up scenario. I’ll say, ‘When I presented it this way in this class, this is how Daniel responded. And they say, ‘Exactly! That’s Daniel. That’s how we need to reach him.’”⁹⁶

Lynn Kuykendall, the South Carolina Master Teacher, believes field-testing garners much higher teacher “buy in” during professional development time. “I have not had one of my Career Teachers out on a cluster meeting day this whole school year, which is amazing,” she says. “Prior to TAP, I can’t tell you, when we would meet at 3:30 to 5 pm at the end of a long work day, how many teachers had appointments and couldn’t stay for professional development. There were hundreds of excuses, and who could blame them? Why would they buy in to something offered by an outsider when that person could offer no evidence it would work with their kids in their classrooms? But now it’s only strategies that we have field tested in this very school with the students they teach. So they know we won’t waste their time.”⁹⁷

TAP’s design also addresses criticisms that most professional development includes little follow up and accountability. Case in point: using funds from a Texas Educator Excellence Grant—a state program that funds pay-for-performance initiatives in high-poverty, high-performing schools—Velez’s school implemented a different compensation reform design prior to adopting TAP. The grant included funds that could be used for professional development. “The professional development was not as structured,” she explains. “The campus did have the flexibility to decide what professional development to choose, but there was no follow-up piece to it that ensured that the teachers were going to take this into their classrooms and truly use it to help improve student achievement.” Now, she says, “we provide the model lessons, team teaching, or time to go observe another teacher, and we give them very detailed feedback for increasing the success of the strategy within their classrooms. Before, we would go to professional development and it would be great for that one day or two days, but then we’d come back and wonder, ‘So how does this fit what I’m doing?’”⁹⁸

“And then there’s the accountability piece of the professional development,” she continues. “We expect teachers to bring something back for the next cluster meeting after using the strategy. Whereas with other professional development, there is nothing that you can really hold teachers accountable for, to bring back or look at. We call it the teachers’ ‘homework,’ and students will ask them if they did their homework.” Vicky Condalary, an Executive Master Teacher with the Louisiana Department of Education, adds that Master Teachers can spot an implementation problem if the quality of the student work teachers bring back does not match what they saw during field testing.⁹⁹

Finally, other activities in cluster group meetings can also create strong vertical alignment. Consider the school improvement plan at Clinton Elementary School in South Carolina. “We take our students and we highlight their names if they need to move more than a year’s growth in a year,” says Master Teacher, Lynn Kuykendall. “We know them by name when we’re meeting in cluster, we call them by name. How is Susie doing? We know she has to grow more than a year this year, so is she on that trajectory and if not what do we need to do about it. We refer to them by name both in our school plan and then in our cluster meetings. Before we’d create a plan and maybe just put it on the shelf, or look at student data and work up a generic plan and put it aside.”¹⁰⁰

Forms of alignment in TAP

The description above illustrates and adds flesh to the forms of alignment identified by the Consortium for Policy Research on Education's Strategic Management of Human Capital project based on ideas from the private sector. TAP reveals two mechanisms for vertically aligning compensation and other HR policies to system goals for student achievement and instructional improvement.

The first is direct alignment to student achievement goals and schoolwide improvement plans. Professional development is directly linked to school improvement through incorporation of field-tested strategies identified by Master Teachers to support the school's specific student achievement goals, and also analysis of student progress toward the goals during common planning time. Compensation is at least indirectly linked to student achievement goals since annual bonuses are based in part on classroom and schoolwide achievement gains. However, bonuses are based on overall gains, not just gains in the specific subskills identified in improvement plans, such as making inferences from text.

The second mechanism for vertical alignment is the adoption of the Teaching Skills, Knowledge and Responsibilities Performance Standards, or "TAP rubric," which allows each HR area to align with a common set of teacher competencies that reflect the skills teachers need to support system goals for student achievement and school improvement. As in the Strategic Management of Human Capital model, each component of HR—evaluation, professional development, compensation—is linked to the rubric, which functions as a set of "job competencies" based on a detailed vision for good teaching.

The rubric also functions as a mechanism for horizontal alignment, enabling compensation, professional development, and evaluation to support and reinforce one another. Because each HR area is linked to the rubric, all of the HR areas support and reinforce one another. Such a system is profoundly different from a traditional system in which those HR functions either are driven by different visions of good teaching, such that they end up sending very mixed signals or, indeed, no clear signals at all. For example, in a typical school, evaluations are based on a "checklist" of teacher characteristics and behaviors that do not relate to the knowledge and skills teachers are asked to learn during professional development time. And compensation is linked to neither, since it only rewards longevity and graduate credits.

In contrast, in a TAP school, compensation in the form of variable pay is based partly on demonstrating core competencies in the TAP Rubric. Formal teacher evaluations are based on the same rubric and provide specific feedback related to it. Professional development, then, is geared toward helping teachers improve on the rubric and learn new instructional strategies that are integrated with it. Teachers who reach high levels of performance on the rubric can advance their careers and their compensation in the form of base pay by assuming responsibility for evaluating and developing other teachers against the rubric.

In addition, the school improvement process feeds field-tested instructional strategies proven to raise student achievement directly into the professional development process, which improves teachers' ability to earn bonus compensation (variable pay) tied to classroom and schoolwide learning gains.

According to the review of performance pay studies by CPRE's Heneman, Milanowski, and Kimball, this kind of alignment is critical for performance pay to be perceived as fair and reasonable by teachers. "Teachers must see an effort-performance link, they note in their report. "Often referred to as teacher expectancy, this link is a subjective one in which teachers judge the probability that a focused, intensive effort on their part will result in the desired performance," necessary to earn the additional compensation. Thus, "the performance pay plan must provide teachers with every possible opportunity to be successful in their performance."¹⁰¹

However, as described above, most schools deny teachers the kind of training, time, and feedback necessary to improve their performance. The solution in TAP is to have teacher-leaders share responsibility with administrators for overseeing very high-quality teacher evaluation and professional development systems. But that is only made possible through base pay augmentations that permit schools to hire Mentor and Master Teachers who can assume those "additional roles and responsibilities."

That suggests a second kind of horizontal alignment which has received very little attention from policymakers—alignment as capacity-building. Indeed, a close examination of TAP's design reveals that the benefits of alignment can extend beyond what the words "support and reinforce" might initially suggest. In the TAP model, compensation is leveraged to build the necessary system capacity to significantly improve other HR functions such as evaluation and professional development.

Here is how it works. First, base pay is augmented (or base pay progression is reformed) to allow schools to "pay for additional roles and responsibilities." Then, those roles and responsibilities are intentionally designed to include shared management of other HR functions such as professional development and teacher evaluation. That, in turn, creates a team of teacher-leaders who have the authority, the expertise, and the time to greatly improve professional development and teacher evaluation.

In short, TAP's design deploys its less famous compensation reform—base-pay augmentation for roles and responsibilities—to greatly boost capacity for instructional leadership in schools.

The distinction between simply aligning HR practices and improving them is an important one. Aligning compensation to professional development and evaluation via a common teaching competencies framework does not necessarily mean schools will dramatically improve both functions. Some districts studied by CPRE researchers, for instance,

had adopted Danielson's Framework in order to tie performance pay to better teacher evaluations. Yet the new evaluation systems suffered because principals did not have the expertise or time to manage a more sophisticated evaluation system. As Shields and Miles of Education Resource Strategies put it, the large so-called "review span" in traditional schools makes it nearly impossible for principals working on their own to conduct thorough and useful teacher evaluations.

In contrast, TAP lowers review span by leveraging "pay for roles and responsibilities" to create a team of teacher-leaders who can be trained to help conduct evaluations. For example, assume that a TAP school with 30 Career Teachers employs two Master Teachers and four Mentor Teachers, meeting the minimum guidelines for 15 Career Teachers per Master and eight Career Teachers per Mentor recommended by The National Institute for Excellence in Teaching. That would provide the school with seven trained evaluators including the principal to evaluate 34 Career and Mentor Teachers who have full- or part-time teaching duties. The review span in such a school would round up to five,¹⁰² much lower than the review span in traditional schools and at the lower end of review spans in leading-edge schools studied by Shields and Miles.

Thus, horizontal alignment in TAP does not just derive from using the same rubric to compensate, evaluate, and provide professional development to teachers. It also comes from purposefully delegating authority for professional development, teacher evaluation, and school improvement to a group of teacher-leaders who are responsible for ensuring that all of those functions are implemented in highly integrated ways. TAP's designers need not necessarily have combined all of those functions into the two job descriptions. The designers, for example, might have followed the advice of most experts on instructional coaching and decided that professional development should not be conducted by the same staff members responsible for formal teacher evaluation.

By designing the career ladder roles the way they did, however, TAP's founders seem to have created a kind of self-reinforcing "alignment loop" between the two kinds of differentiated pay in the model. The improved and aligned versions of professional development and teacher evaluation made possible by base-pay augmentations in turn support and reinforce the annual performance bonuses by providing teachers with high-quality opportunities to develop the very skills rewarded by the bonuses.

Moreover, new teachers over time will be given the support they need to reach high enough levels on the competency framework to become eligible to become Mentor and Master Teachers, thereby earning higher base pay. Those teachers will then provide high-quality professional development and evaluations that enable other teachers earn annual bonuses, and so on.

Such an approach stands in sharp contrast to the very fragmented HR systems (if they can even be called that) in traditional schools, where compensation is disconnected from teacher evaluations, which in turn are disconnected from professional development and

which often is only loosely connected to school improvement planning. For teachers in most such schools, evaluation “happens,” professional development “happens,” getting paid “happens,” and school improvement plans get written that might or might not cause other things to “happen.” But none of those experiences has much to do with any of the others. And each demands and rewards a different and unrelated set of behaviors.

In fact, it might be more accurate to call the TAP design an “integrated system” rather than just an “aligned design” or a “comprehensive model.” For teachers working in a TAP school, it is difficult to tell where professional development, teacher evaluation, and school improvement begin and end because they are all part of one fluid process. Indeed, in a case study of the Minnesota Q-Comp program inspired by TAP, Heneman points out that such a design should have the “side effect” of improving talent acquisition as well as talent development: “Such a transformed workplace itself can become a powerful recruitment tool for districts,” he explains in his report.¹⁰³ Other research suggests that a workplace characterized by high levels of instructional leadership and support also can reduce attrition among beginning teachers.¹⁰⁴

Of course, because this analysis is based on only one school-level model, it is important to keep in mind that, from a larger system perspective, there are many HR areas beyond professional development and evaluation to consider. The following examples, one state-level and the other district-level, illustrate the importance of aligning compensation with recruitment and selection:

- From 2001 to 2004, North Carolina provided a recruitment bonus of \$1,800 for certified math, science, and special education teachers to work in high-poverty or low-performing schools. An analysis of the effects showed that while the program reduced turnover in such schools by about 17 percent, the effect could have been much larger if the state had better communicated to teachers about criteria for the bonuses.¹⁰⁵
- Studies by the New Teacher Project have found that recruitment, selection, and placement policies in urban school systems create significant obstacles for strong applicants to be hired to work in high-need districts and schools.¹⁰⁶ Such hiring practices would create profound misalignment in states or districts that decided to pay large recruitment bonuses for strong teachers to work in “hard-to-staff subjects” or “hard-to-staff schools” in urban districts.

Finally, TAP’s aligned design illustrates that it is important to think of alignment as a two-way street, especially when it comes to compensation. Better teacher evaluations and professional development support and reinforce differentiated compensation, and differentiated compensation creates the instructional leadership capacity to implement better teacher evaluations and professional development. Therefore, it is helpful to think of alignment through two distinct lenses that, for lack of better terms, might be called “front end” and “back end.”

When assessing the alignment between teacher evaluation and differential compensation, the key questions would be the following:

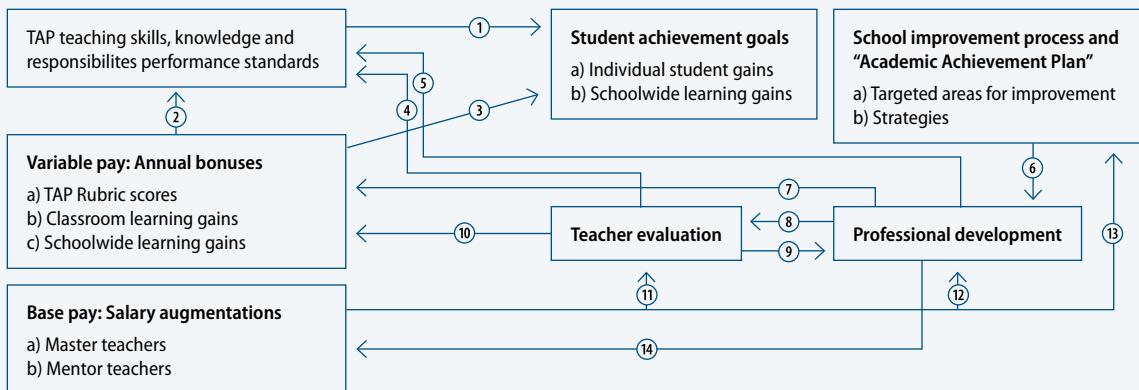
- Front end alignment: Will our teacher evaluation practices support and reinforce the differentiated compensation model we are adopting, and, if not, how can we ensure that they will?
- Back end alignment: Will the differentiated compensation model we are adopting support, reinforce, and—ideally—create capacity to improve our teacher evaluation practices?

Taking all of the above into consideration, assessing horizontal alignment of compensation involves at least four questions each about recruitment, selection, induction, mentoring, evaluation and performance management, and professional development. Those four questions are:

- Are Policy A and Policy B encouraging acquisition and development of the same teacher competencies?
- Are Policy A and Policy B misaligned in such a way that one, or both, subverts the other?
- How can Policy A be specifically designed to support Policy B by creating more capacity for doing Policy B well?
- How can Policy B be specifically designed to support Policy A by creating more capacity for doing Policy A well?

Table 1 in the “Recommendations” section below provides a framework for policymakers to use when assessing proposed compensation reforms through all of these lenses. Before turning to that table, however, this paper will first examine how to build more teacher capacity in our school systems by paying teachers to take on additional roles and responsibilities.

Forms of alignment in the TAP model



1. Research has demonstrated a correlation between teachers' scores on the TAP Rubric and classroom learning gains, and the School Leadership Team cross-checks both sources of data for validity.
2. Annual bonuses are based in part on teachers' scores on the TAP Rubric.
3. Annual bonuses are based in part on classroom and schoolwide learning gains.
4. Teachers are evaluated by the principal, Master Teachers, and Mentor Teachers who conduct four to six classroom observations and score teachers' instruction against the TAP Rubric.
5. Cluster meetings and individual coaching help teachers improve their performance on the TAP Rubric.
6. Based on analysis of student data, Master Teachers lead an effort to identify areas for improvement (e.g., reading comprehension); create a school Academic Achievement Plan; identify instructional strategies to address the areas identified for improvement; field-test and fine-tune the strategies with students in the school; and train teachers to use the strategies through professional development activities (weekly cluster meetings, individual coaching).
7. Professional development helps teachers improve their scores on the TAP Rubric and provides them with field-tested strategies proven to produce learning gains for students in the school.
8. After each classroom observation, teachers receive coaching to help them improve on an "area for refinement."
9. Scores from teacher evaluations are fed into the CODE data system, which Master Teachers can analyze to identify areas of the TAP Rubric for special attention in professional development (cluster meetings, coaching).
10. Results of teacher evaluations account for a portion of annual bonuses, and post-conference sessions with the evaluator provide teachers with specific feedback including an "area for reinforcement" and an "area for refinement" on the TAP rubric, along with follow-up coaching.
11. Base pay augmentations allow the school to promote or hire a team of Master Teachers and Mentor Teachers who boost the school's capacity to implement a more sophisticated and rigorous teacher evaluation system that requires four to six observations and post-conferences for every teacher every year.
12. Base pay augmentations allow the school to promote or hire a team of Master Teachers and Mentor Teachers who boost the school's capacity to implement ongoing, intensive professional development through weekly cluster meetings, team teaching, demonstration lessons, modeling of Rubric skills and new instructional strategies, and other kinds of support.
13. Base pay augmentations allow the school to promote or hire Master Teachers who boost the school's capacity to conduct a more sophisticated school improvement planning process—analyzing student data, identifying areas for improvement, and field-testing new instructional strategies to address those areas.
14. Over time, evaluations and professional development should help teachers develop high levels of expertise on the TAP Rubric and effectiveness in the classroom, which, along with other skills, will qualify them to apply for Mentor Teacher and Master Teacher positions and earn base pay augmentations.

Building leadership capacity by paying for additional roles and responsibilities

So far, the discussion about compensation reform has focused much more on performance pay than on pay for roles and responsibilities. Moreover, much of the attention on alignment has focused on how other HR policies such as teacher evaluation and professional development can be used to build capacity for performance pay. To some extent this is because those additional HR functions build what might be called “political capacity” for compensation reform. Basing performance bonuses on teacher evaluations rather than only on student achievement gains creates more “buy in” from teachers and teacher organizations. Providing funds for professional development sweetens the pot, too.

However, the analysis of TAP above suggests that the model’s less-discussed vehicle for compensation reform, pay for additional roles and responsibilities, might be the more powerful engine for comprehensive change since it can be leveraged to dramatically improve a host of other school practices, including professional development and evaluation. But TAP evidence also suggests that those benefits only accrue when sufficient attention is given to designing roles that focus on instructional improvement and empower teachers to exercise true leadership in their schools.

That raises a set of questions that deserve further exploration. First, how many teachers and potential teachers are interested in taking on such roles and responsibilities? Second, what are the challenges to designing roles and responsibilities that focus on instructional leadership, with an emphasis on “leadership”?

When answering these questions we must also consider how more pay for additional roles and responsibilities can be an important tool for compensation reform. Evidence suggests that younger adults and subject-area specialists would be more likely to enter and remain in teaching if it offered better career advancement opportunities. Research conducted by Harvard University’s Project on the Next Generation of Teachers found that, unlike their predecessors, many beginning teachers (including mid-career entrants as well as recent college graduates) have a strong interest in differentiated, “hybrid” roles that would allow them to continue teaching while moving beyond the classroom to have a greater influence in their schools.¹⁰⁷

Other research suggests the “flat,” undifferentiated career structure of teaching makes it difficult to attract and retain talented college graduates and subject-area specialists. A national poll by Public Agenda found that almost seven in ten young college graduates

(69 percent) think that teaching does not offer good opportunities for advancement.¹⁰⁸ And among math teachers who left their schools in 2000-2001, nearly half (46 percent) cited inadequate opportunities for professional advancement as a reason for doing so—the fourth-ranked reason out of 13 choices.¹⁰⁹

Beyond TAP, what do large-scale efforts to pay for roles and responsibilities suggest about designing such opportunities? Unfortunately, there is little evidence to examine. Following the publication of *A Nation at Risk* in 1983, many state policymakers became interested in “career-ladder” programs, which were incentive plans that enabled teachers to earn higher salaries by demonstrating good performance or taking on more responsibilities. By the fifth anniversary of the report, 14 states had adopted career-ladder programs, eight were developing them, and 17 were conducting smaller scale initiatives and pilot programs in some districts.¹¹⁰ Most of those programs evaporated due to funding problems and concerns about the fairness of criteria used for advancement. By 1994, only four states still funded career-ladder programs.¹¹¹

Two of those programs still exist—Arizona’s career-ladder program, which dates to 1984, and Missouri’s, which dates to 1987. Both provide a strong contrast to TAP’s approach. Both states offer broad flexibility to districts in structuring roles and responsibilities, and district plans generally allow teachers to accumulate sets of additional responsibilities from a wide menu of options rather than moving into defined, differentiated positions as in TAP. At the risk of oversimplifying, it is useful to think of a “role” as something akin to a workplace position, a set of focused and coherent duties that can be described in a job description. In contrast, a “responsibility” can be thought of as an add-on obligation or activity that does not change the formal job description itself.

For example, the Missouri State Board of Education requires only that “each career ladder stage shall contain responsibilities commensurate and adjustable to the compensation offered for that stage that will be completed by the teacher while on the career ladder,” and that the responsibilities should “obviously relate” to formal school improvement, curriculum development, professional development plans, or “instructional improvement.”¹¹² At each stage of the ladder, teachers must log a certain number of hours engaging in appropriate activities. However, according to data collected by the Missouri State Department of Education, in the 2007-8 school year about half of the over 2.2 million hours logged by career-ladder teachers were spent on activities that are not about exercising instructional leadership, such as “parent contact,” “student tutoring,” and “other student contact.”¹¹³

Arizona’s law requires only that career ladder plans include “opportunities for higher level instructional responsibilities.”¹¹⁴ Older career-ladder laws seem to have spent much more ink defining how teachers could move up the ladder than describing what kinds of work they would contribute as they climbed. Arizona, in contrast, grants wide latitude to districts in defining such responsibilities, and, as a result, districts vary widely in their designs.

In nearly every case, districts allow teachers to select from a menu of activities, or “responsibilities,” rather than offering them differentiated positions or “roles.” Some districts require the activities to exhibit greater levels of leadership as teachers move up the ladder. In other cases, districts pass on the flexibility granted by the state and permit teachers wide latitude to propose their own activities for meeting the requirement.¹¹⁵

Some Arizona career-ladder districts strongly encourage teachers to focus their time on activities related to instructional leadership as they move into the higher rungs. Amphitheater Public Schools in Tucson requires teachers reaching the highest rung of the ladder to engage in “collaborative action research” to identify strategies for improving student achievement for several years. After that, teachers may choose from a short menu of activities:

- Engage in a “full review” of their practice involving peer and administrator evaluation and self-reflection
- Conduct independent action research
- Mentor younger teachers
- Provide professional development to other teachers based on an analysis of student data.¹¹⁶

Other Arizona career-ladder districts allow teachers on even the highest rungs to meet the “higher level instructional responsibilities” by spending at least some of the time engaged in activities that are neither related to instruction nor require much leadership. A number of districts permit highest-rung teachers to fulfill some or all of the requirement through activities ranging from editing parent newsletters to advising extra-curricular clubs to coordinating co-curricular events such as spelling bees.¹¹⁷

To be clear, this is not meant to be an analysis of the programs’ impact on student achievement, but rather only an examination of how they define the “roles and responsibilities” elements of their career ladders. A 2007 evaluation of Arizona’s program found that on average, students in career-ladder schools perform better on state tests than students in non-career ladder schools, after adjusting for differences in student and school characteristics.¹¹⁸ A 2008 study of Missouri’s program found some evidence of a small positive impact on math, but not reading, achievement.¹¹⁹

Since it is based on TAP, Minnesota’s Q-Comp program offers a useful contrast to the Arizona and Missouri programs. Q-Comp requires districts to create a career ladder or multiple career paths for teachers that result in very clearly defined roles or positions. The Minnesota Department of Education grant application form requires applicants to provide, for each role:

- The position title
- The job description
- The qualifications
- The responsibilities

- The hiring process
- The evaluation procedures
- Any additional compensation.

In addition, each applicant must provide a narrative summary of how the positions will include duties related to “field-testing and researching instructional strategies, providing professional development, coaching, observation and evaluation, and mentoring.”¹²⁰

According to Beth Driscoll, who coordinates Arizona’s career ladder program at the State Department of Education, “the higher level responsibilities are probably the most challenging aspect for the districts, to first of all have enough opportunities for all the teachers who participate, and second of all make them meaningful and valuable and have the highest impact for all the levels on the ladder.” Yet she believes the flexibility of the program is part of its appeal.

Moreover, she says, a design that provides only higher-status roles might not appeal to some teachers because it conflicts with the strongly egalitarian nature of the teaching profession.¹²¹ Broadly speaking, then, policymakers seem to face at least three major sets of questions when designing pay for additional roles and responsibilities:

- Should career advancement designs permit more flexibility and allow teachers to select from a broad menu of allowable responsibilities, as in some current state career ladder programs? Or should they create more coherent roles and positions as in Minnesota’s TAP-based Q-Comp program?
- How should policymakers deal with the issue of professional status? Should they sidestep it in hopes of garnering greater initial teacher buy-in? Or should they intentionally create career pathways that might appeal to teachers who would value status-enhancing advancement opportunities?
- How should the design deal with issues of access? Should the ladder narrow as teachers climb it? Or should systems offer a variety of “hybrid” roles for many teachers?

Figure 3 offers a visual representation of the first two issues: coherent roles vs. narrower responsibilities and status-neutral vs. status-enhancing opportunities. The lower left quadrant represents a “career-enhancement” design or, more simply, “pay-enhancement” design, while the upper right could be said to represent a true “career-advancement” approach.

FIGURE 3
Different Ways to Design “Additional Roles and Responsibilities”

	Responsibilities	Roles
Status-Enhancing	Chairing or serving on a district committee	Master teacher
Status-Neutral	<ul style="list-style-type: none"> Tutoring students Providing parent workshops Sponsoring an extra-curricular club or co-curricular activity 	On-track monitor

Regarding the third question (How should compensation for additional roles and responsibilities deal with issues of access to such opportunities?) Kristan Van Hook, NIET's senior vice president for public policy and development, cautions that it is important to distinguish between performance pay and career pathways in newer designs like TAP. "We can recognize high performance in many ways," she says. "But those instructional leadership roles are not simply rewards for high performance *per se*. They're specific jobs in the school that have to be filled by people with very particular kinds of expertise."¹²²

In other words, TAP's "Multiple Career Paths" component has two distinct goals—offering career advancement opportunities and creating much-needed capacity for instructional leadership—with the latter being just as important as the former.

At the same time, Van Hook says, "there are no limits on who can receive performance bonuses or other kinds of recognition for excellent work in the classroom." In fact, under its federal Teacher Incentive Fund grant, South Carolina's TAP program has created an additional way to recognize consistently high-performing Career Teachers through a kind of super bonus. Teachers whose students achieve higher than average growth for multiple years in a row would become eligible for especially high performance bonuses over time.

TAP also provides evidence for thinking about how many instructional leadership roles policymakers might need to create. Given the need to build much stronger instructional leadership in high-needs schools that have adopted TAP, NIET recommends that schools aim for a ratio of 12 to 15 Career Teachers per Master Teacher and 6 to 8 Career Teachers per Mentor Teacher. That translates into roughly 20 to 25 percent of a school's faculty taking on such leadership roles at any given time. NIET contends that such a design provides a great deal of opportunity for advancement.

Finally, many districts in Minnesota receiving Q-Comp funds have designed more than two career-ladder roles focused on instructional leadership. Therefore, the issue of access might better be viewed as a much longer-term issue rather than an immediate problem facing policymakers now.

Moreover, if access ever became a major problem, policymakers could create additional roles. School systems, for instance, might offer roles focused on student support, perhaps building on evidence of what works to improve achievement or graduation rates in high-performing places or in research-proven programs. Such roles might appeal to classroom teachers who want to take on greater challenges but are uncomfortable with status issues raised by training or supervising other adults. For example:

- In top-performing Finland, every school employs "special-needs teachers" who receive additional training to provide individual or small-group support to students who need it based on their academic progress, mainly in language arts and mathematics. On average, about 30 percent of Finnish students receive such additional help every year, including some "strong students" who are struggling to master a particular topic or skill.¹²³

- Of the four dropout prevention programs in the United States identified as having the strongest proven effects by the federal government’s What Works Clearinghouse, two programs relied on adults who were trained to closely monitor individual student data for early warning signs that predict dropping out, then to immediately intervene to help students solve whatever is causing the problem, and then to provide general forms of advocacy and guidance to small groups of students identified to be at greater risk of dropping out. One such program, the University of Minnesota’s Check & Connect model, which calls the position “Monitor,” has been implemented successfully in multiple places.¹²⁴ To ensure this is a role rather than an add-on responsibility, systems might create teacher-leader positions that focus broadly on raising graduation rates.

Such roles would not create as much integration and “horizontal alignment” as instructional leadership positions such as Master Teacher and Mentor Teacher, but both would be “vertically aligned” given federal and state accountability policies focused on closing achievement gaps and improving graduation rates. Again, however, given the tremendous gap between demand for better professional development and evaluation and the amount of leadership capacity currently available, policymakers might prioritize instructional leadership roles first.

The issue of professional status presents a different challenge. A recent study by a team of researchers affiliated with Harvard University’s Project on the Next Generation of Teachers interviewed third to tenth year teachers taking on instructional leadership roles and found that traditional norms of the profession presented teacher-leaders with significant obstacles. Teachers who took roles requiring them only to support their colleagues’ classroom practices encountered no resistance from peers. However, “reform roles” intended to change colleagues’ practices engendered considerable resentment and sometimes outright resistance.¹²⁵

Such roles “introduce a triple threat to the traditional norms of the profession—egalitarianism, autonomy, and seniority,” the researchers concluded. “Our findings suggest that the traditional norms of teaching remain potent, and that current reformers and school leaders who seek to increase the instructional capacity of their schools must take that into account when they design and introduce differentiated roles for teachers.”

The same study found that teachers taking on “reform roles” had to figure them out in isolation without support from other teacher-leaders or administrators, which sometimes caused them to lower their expectations for change. The researchers concluded that leadership roles might be necessary to attract and retain talent for the next generation of teachers as well as increase instructional leadership in schools, but that policymakers and administrators must provide teacher-leaders with a great deal more support to help them negotiate such roles.”¹²⁶

Some experts believe that the traditional norms of egalitarianism and privacy represent a significant obstacle to school improvement and must be overcome. “One of the strongest social norms among school faculty is that everyone is expected to pretend that they are equally effective at what they do,” observes Richard Elmore, a professor of educational

leadership at Harvard University. “Yet the entire process of improvement depends on schools making public and authoritative distinctions among teachers and administrators based on quality, competence, expertise, and performance. If everyone is equally good at what they do, then no one has anything to teach anyone else about how to do it better.”¹²⁷

The status issue comes into particular focus when deciding whether teacher-leaders should have responsibility not just for supporting their peers but also for evaluating them. To complicate matters, most proponents of coaching and mentoring strongly argue that coaching should be kept strictly separate from formal evaluation, and some experts even caution against providing any kind of evaluative feedback whatsoever. The assumption that coaching must be “non-evaluative” is predicated on the idea that teachers must be able to trust their coaches to confide in them, and it can be traced back to some of the earliest articles on the topic. Beverly Showers, a founder of the instructional coaching movement, claimed in 1985 that “the norms of coaching and evaluation practice are antithetical and should be separated in our thinking as well as in practice.”¹²⁸

No formal research study has been conducted on the success of giving TAP’s Master and Mentor Teachers responsibility for evaluation along with their instructional coaching duties. But a study of a peer assistance and review program in California confirms TAP’s anecdotal evidence that, under the right conditions, teacher-leaders are capable of enacting roles that combine support with evaluation. Analyzing a peer assistance and review program in a mid-sized urban district in California, Jennifer Goldstein, an associate professor in the Baruch College School of Public Affairs of the City University of New York, found that, while trust was indeed an important factor in the assistance mentees received and their plans for remaining in the profession, the program actually seemed to foster trust rather than diminish it.

Even though their coaches were also their summative evaluators, mentees reported an overall high level of trust in coaches, Goldstein discovered. Mentees who did not report a high level of trust were low-performing teachers.¹²⁹

In addition, given the significant costs associated with instructional coaches, it seems important for policymakers to consider findings from recent research on “non-evaluative” models of coaching. For example, the researchers studying NSF urban initiative sites concluded that:

*Providing powerful feedback requires that teacher-leaders move beyond mere sharing of expertise and toward a more complicated engagement around the work. In effect, it requires teacher-leaders to challenge teachers to change their practice. Yet, because most teacher-leaders lacked the authority to evaluate teacher performance, or to directly request changes in classroom practice, they delivered feedback that was suggestive, not directive ... Consequently, there is a notable gap (perhaps even a chasm) between the work that teacher-leaders are able (and willing) to do and the expectations for systemic change in instructional practices.*¹³⁰

Another study of a coaching initiative in a large urban district published in the same volume last year found similar problems. “Several coaches in our study expressed concerns about their limited impact because they felt they had to wait for a teacher invitation to visit classrooms or could only make suggestions to teachers about their instruction after observations,” said the report. “When teachers did not incorporate the suggested changes, coaches felt that informing principals might be seen as evaluative.”¹³¹

In fact, the editors of the volume in which both studies appeared concluded that further research was necessary to investigate the assumption that teacher-leader roles must be non-evaluative.¹³²

It might be even be argued that weak HR practices such as inflated evaluations have combined with persistent norms such as egalitarianism and privacy to effectively de-professionalize teachers. Goldstein cites research suggesting that professionalism requires taking collective responsibility for client welfare, advancing the knowledge base, and dealing with weak performers. “Self-regulation, central to professionalism and professionalization, has been slow to occur in education,” she argues. “Policy makers, practicing educators, and the public tend not to believe that teachers are capable of regulating themselves. Yet, with the peer assistance and review program in Rosemont, teachers’ work expanded to include quality control and gatekeeping, with many benefits to the quality of the evaluation process.”¹³³

In addition, there is evidence that education systems in other countries have found ways to provide teachers with career advancement opportunities that directly confront issues of leadership and status. Teachers in Singapore, for example, can choose from among three career tracks—teaching, leadership, and content specialist—each of which has a career ladder based on performance. Excellent teachers can become Senior Teachers and a very elite group can become Master Teachers. “Selectivity operates at every phase of the teacher’s career: the system is always seeking to identify excellence,” observes Susan Sclafani, who wrote about the system in a recent report for the Aspen Institute.¹³⁴

Singapore’s career development and advancement opportunities are supported by a sophisticated Enhanced Performance Management System. “From the first year on campus, every teacher is planning a career and using self assessment, coaching, and evaluation to achieve next steps,” says Sclafani. “Through the EPMS process, teachers are encouraged to expand their teaching repertoire, select a career track, and take those developmental actions that lead to greater competence-and higher levels on the career ladder.”¹³⁵

Singapore’s approach also illustrates the importance of designing HR policies that align on the “back end” as well as the “front end.” On the front end, the great career development and advancement opportunities help the city-state’s Ministry of Education recruit teachers from among the top 30 percent of graduating students.¹³⁶ On the back end, career pathways build significant capacity for instructional leadership. “Reading through this extensive [EPMS] process, one might ask where reporting officers get the time to do all

this discussion, counseling, support, and evaluation,” says Sclafani. “The answer is the distributed leadership model in place. It is not just the principal and vice principal(s) who are responsible for supporting and evaluating teachers.”¹³⁷

Finally, there might be significant cost implications for systems that fail to consider how to align compensation reforms with other human resource strategies, particularly when they fail to consider how differentiated compensation can pay for instructional leadership. Consider a school district that spends considerable amounts on the following functions:

- **Professional development.** Several million dollars to hire instructional coaches to improve math and literacy scores
- **Evaluation and instructional leadership.** Several million dollars to expand schools’ administrative staffs to add co-principals who can provide greater instructional leadership and help conduct more sophisticated teacher evaluations
- **Compensation.** Several million dollars from a state career-ladder grant to pay teachers for “additional roles and responsibilities” that do not support the coaching program or the new teacher evaluation system.

The district might have thoughtful reasons for such decisions. Yet leaders might be wasting millions of dollars by failing to align compensation reform with efforts to improve professional development, teacher evaluation, and instructional leadership. TAP illustrates that in a school using an integrated design for compensation reform, positions such as Mentor Teacher and Master Teacher can provide the capacity to pursue all of those goals at much lower cost.

“The Milken Foundation estimates that TAP costs about \$500 per student,” says CPRE’s Odden. “But if instructional coaches are already part of your basic school aid formula, then TAP would cost much less than that.”¹³⁸

Recommendations

Considering the increasing federal, state, and local investment in performance-based compensation for teachers, it makes sense to ask whether such reforms are getting the “biggest bang for the buck.” This paper has shown that policymakers should ask whether proposed compensation policies will be supported or undermined by other policies in schools and districts. Just as important, they should ask whether new ways of paying teachers—particularly for taking on additional roles and responsibilities—will help realize much-needed improvements in other human resources practices in education, such as teacher evaluation and professional development.

Before federal funding for the Teacher Incentive Fund doubles and potentially quintuples over the next few years, national policymakers should take the time to ask the same questions of that program. Finally, federal policymakers should set an example by funding more research and technical assistance on a potentially powerful yet often ignored kind of compensation reform, paying teachers to assume new kinds of roles and responsibilities in their districts and schools. Policymakers must begin to see compensation reform as a strategy not only for motivating and rewarding individual teachers, but also for building the capacity of our public schools to take on the hard work of systemic improvement so critical for raising student achievement and closing learning gaps between groups.

Let’s consider each recommendation in turn.

When designing compensation reforms, policymakers should consider ways to encourage local school systems to work toward vertical alignment with system goals and horizontal alignment with other HR areas.

This recommendation requires two clarifications. First, the point is not that alignment is strictly necessary to derive any benefit from compensation reform. Indeed, studies have shown that simpler performance-pay initiatives can raise student achievement.¹³⁹ The problem is that such initiatives tend to be short-lived because they fail to garner sufficient buy-in from teachers or from teacher unions. Rather, the point is that alignment can add value to compensation reforms and allow compensation reforms to add value to other HR areas. Second, TAP is not the only way to differentiate compensation or to align compensation reforms with other HR areas. Rather, TAP is presented as a useful illustration and “benchmark” that policymakers can consider as they assess alignment.

In addition, although this paper focused on HR alignment, policymakers might also consider how well compensation reforms will align with other broad system functions beyond HR practices. For example, how will compensation reforms support and reinforce school accountability policies, and vice versa? How can data systems be leveraged to support performance-based compensation? And how can data from compensation reform initiatives be used to boost the power of those systems?

Finally, as illustrated in this paper, alignment is a two-way street. Therefore, when designing compensation reforms, policymakers should consider alignment through the “front end” and “back end” lenses:

- Front-end alignment refers to how other policies and practices can be leveraged to support and reinforce the intended compensation strategy.
- Back-end alignment refers to how the compensation strategy can be leveraged to support, reinforce, and improve other policies and practices.

Table 1 on the following page provides a template for asking critical questions about alignment when designing initiatives to reform teacher compensation.

TABLE 1
Asking the right questions about alignment

		“Front end alignment”	“Back end alignment”
	Will the compensation reform reward the same set of job-related competencies? Are all HR functions based on a common set of competencies?	Will current policies and practices in this area support or subvert the intended compensation strategy?	How can current policies and practices in this area be revised to better support the intended compensation strategy?
HR policies			
Recruitment			
Selection			
Placement and distribution			
Induction and mentoring			
Professional development			
Evaluation and performance management			
Instructional leadership*			
Career progression*			
Leadership development*			
Other policies			
Improvement Goals, plans, and processes			
Accountability			
Budgeting			
Information management			
Student support			

* CPRE’s SMHC Project does not always include these areas in its lists of discrete HR functions. However, they are included here so they are not overlooked and because of their particular relationship with compensation reform.

Policymakers should consider using pay for additional roles and responsibilities to boost instructional leadership capacity in order to improve professional development and teacher evaluation.

This paper has shown that it is important to remember that the benefits of aligning compensation flow both ways. Not only should other policies support compensation reform; compensation reform should be designed to help improve other policies. In particular, one kind of compensation reform, pay for additional roles and responsibilities, can be leveraged to significantly boost instructional leadership and improve teacher evaluation and professional development—but only if it is carefully designed to do so.

To be sure, simply directing base-pay progression away from strict reliance on criteria such as experience and education credits and toward actual expertise and effectiveness will deliver benefits in and of itself. Those dollars will offer incentives that are much more vertically aligned with the system’s goals for student achievement and instructional improvement, and by allowing excellent teachers to earn more money earlier in their careers, they will help attract and retain talent in the profession. But if those redirected dollars also pay for teachers to engage in school leadership roles that focus on improving instruction, then policymakers can build significant capacity to improve other HR practices such as teacher evaluation and professional development.

Obviously, programs that fund only variable bonus pay tied to student outcomes will have a difficult time building instructional leadership capacity on the back end. But any policy that addresses base pay progression could promote back-end alignment. For example, any policy that requires or encourages districts and schools to create a career ladder, offer career advancement opportunities by offering additional pay for additional responsibilities, reforming the steps and lanes of the salary schedule, or offering “salary augmentations” in addition to annual performance bonuses could encourage back-end alignment.

Of course, teachers who demonstrate advanced expertise and effectiveness but who wish to focus on classroom teaching should be able to earn higher pay on a faster track, too. However, because of the strong interest in career advancement among many young people—along with the huge gap in instructional leadership in most education systems—policymakers should strongly consider offering additional pay for those who also are qualified and willing to take on leadership roles.

With that in mind, policymakers might encourage back-end alignment in any number of ways. For example, they could:

- Require districts or schools to offer advanced roles rather than just permitting accumulation of additional responsibilities, and then define those roles from the top. That might work for initiatives targeted toward smaller numbers of districts or schools. Then

Sen. Barack Obama’s “Innovation Districts” bill, which would have provided grants to 20 school districts, would have required a standard career ladder with steps and defined roles for Novice Teachers, Career Teachers, Mentor Teachers, and Master Teachers.

- Require roles, but permit districts some flexibility in designing such positions. In designing initiatives meant to have broad impact across districts, federal and state policymakers might need to find a “tight-loose balance” that encourages better-aligned designs while offering sufficient local flexibility. Minnesota’s Q-Comp application, for example, ensures that grantees design coherent roles focused primarily on instructional leadership, which should in turn build capacity to improve professional development and teacher evaluation. At the same time, Q-Comp provides flexibility to grantees in the number of roles and type of roles, allowing significant variation across systems. Forest Lake Area Schools in Minnesota designed four new roles: Instructional Coach, Mentor, Direct Instructional Leader, and Building Instructional Leader. Orono Public Schools designed five: Professional Learning Community Team Leader, Professional Learning Community Building Leader, Teacher Coach, Peer Evaluator, and Q-Comp Teacher Leader.
- Require grant applicants to answer very specific questions about how they designed the roles and responsibilities in their plans, with an eye to encouraging them to consider a range of alignment questions. For competitive grants, state or federal policymakers could award points based on the responses.

For example, grant applicants might be asked to answer the following kinds of questions:

- How will the new roles and responsibilities support the district or school’s specific goals for student outcomes and the district’s specific educational improvement strategy?
- How will the new roles and responsibilities help your school or district improve other core functions, such as teacher evaluation, professional development, new teacher induction and mentoring, and school improvement planning?
- How will the new roles and responsibilities help attract talented teachers to the school, to the district, or to particular high-need schools in the district? How will the new roles and responsibilities retain talented teachers in the school, in the district, or in particular high-need schools in the district?
- How will the new roles and responsibilities help teachers improve their classroom performance (knowledge, skills, and effectiveness)? Consider not just teachers who take on the roles and responsibilities, but also the novice teachers and veteran teachers who have not yet advanced.

Federal policymakers should carefully review the Teacher Incentive Fund to ensure that it encourages applicants to design performance-based compensation plans that align and improve (rather than just include) professional development and teacher evaluation.

In terms of “front-end” alignment, TIF authorizing language tied compensation to evaluation by requiring that performance-based compensation systems consider “classroom evaluations conducted multiple times during each school year” in addition to student achievement gains. And, although that language did not mention professional development, the U.S. Department of Education added professional development to the selection criteria in the Federal Register notice inviting TIF applications. However, it is not clear that the department considered whether applicants described a specific method for aligning such elements during the application review process.

Moreover, so far the department seems to have completely ignored capacity-building “back-end” alignment even though the extremely brief, 175-word authorizing language very prominently requires that performance-based compensation “provide educators with incentives to take on additional responsibilities and leadership roles.” For example, the department includes only two very cursory mentions of that requirement in its official 4,300-word FAQ document on the TIF program.¹⁴⁰ And, although the *Federal Register* notice inviting TIF applications mentions pay for additional responsibilities and leadership roles once as a requirement of the authorizing language, it includes nothing related to that requirement in either the formal “selection criteria” or the “competitive priorities.”¹⁴¹ Clearly, up until now, this aspect of the TIF program has been left entirely to the discretion of local grantees.

In May, the department’s fiscal year 2010 budget proposal signaled an intention to better address alignment issues in TIF: “Beginning with the competition that the Department will conduct this year with the \$200 million in additional funds provided for this program by the Recovery Act, the Department will place a priority on the support of comprehensive, aligned approaches that support improved teacher and principal effectiveness and help ensure an equitable distribution of effective educators”¹⁴²

To better encourage alignment and integration, the department first should ensure that the required national evaluation of TIF specifically examines the roles and responsibilities question to assess the range of design decisions among local grantees and the extent to which grantees have considered back-end alignment. What kinds of roles and responsibilities are teachers being paid to undertake? Are those roles and responsibilities aligned vertically with grantees’ goals for student outcomes and improvement plans? Are they focused on providing instructional leadership to build capacity for higher-quality and better-aligned PD and teacher evaluations? Or do they resemble the wide range of additional responsibilities allowable under the Missouri and Arizona career ladder programs?

Secondly, the department should revise the selection criteria and competitive priorities to encourage better alignment of teacher evaluation and professional development with

performance-based compensation, and vice versa. This should be done in a way that specifically encourages both front- and back-end alignment between compensation reform and other policies.

Front-end alignment. How will teacher evaluation and professional development support and reinforce performance-based compensation? It is not clear that the department can or should require grantees to adopt a common teacher competency framework as a specific mechanism for horizontal alignment. However, it can require grantees to adopt some mechanism for alignment or at least explain how they plan to encourage it. Through the process, applicants should be encouraged to consider questions such as the following:

- How will teacher evaluations be realigned to measure the same instructional skills rewarded by performance-based compensation?
- Will the evaluations assess and provide feedback on instructional skills known to produce the kinds of achievement gains rewarded by the performance-based compensation?
- How will professional development provide teachers with support to develop the skills measured by the evaluations and strategies that have been shown to produce the kind of learning gains rewarded by performance-based compensation?

Again, it is not clear that the department can require grantees to include “field-tested” strategies as in TAP. However, the design should allow teachers to see a clear link between what they are learning through professional development and the student outcomes rewarded by performance-based compensation.

Back-end alignment. How will the performance-based compensation, including “incentives to take on additional responsibilities and leadership roles,” create opportunities to promote better professional development and teacher evaluations? Again, it is not clear that the department itself should define the specific roles and responsibilities grantees must include in their designs. However, the department must find a better “tight-loose” approach that encourages greater alignment while offering sufficient local flexibility in program design. Applicants should be encouraged to consider questions such as the following:

- Does the design include true “leadership roles” as well as just “additional responsibilities”?
- How will those new roles build capacity for instructional leadership to improve classroom teaching?
- Will the leadership roles build the system’s capacity to provide higher-quality and better-aligned professional development and teacher evaluations?
- If the leadership roles do not include responsibility for helping principals evaluate teachers, how will the system attain a “span of review” that provides enough evaluators to implement “classroom evaluations conducted multiple times during each school year” as required by the TIF authorizing language?
- If the roles and responsibilities focus exclusively on supporting, developing, or coaching teachers, will those individuals have enough authority to provide the “hard feedback” teachers need to improve their practice?

The U.S. Department of Education and others should sponsor more research and provide better technical assistance on compensation reforms that pay teachers for additional roles and responsibilities.

So far such research has focused much more on the three other kinds of compensation reform—market-based incentives, paying for knowledge and skills, and performance pay. If extra pay for additional roles and responsibilities is going to be part of teacher compensation reforms, then policymakers will need more information on the tradeoffs between different designs and much better guidance on issues related to HR alignment.

For example, the department could provide additional technical assistance on pay for roles and responsibilities through the federally funded Center for Educator Compensation Reform. The center “serves as the primary online repository for information, tools, and resources to support Teacher Incentive Fund grantees, policymakers, state officials, and district professionals with the design and implementation of educator compensation reform programs and policies.”¹⁴³ However, although it offers useful resources on many compensation-related topics, the center offers no guidance on how to structure pay for additional leadership roles and responsibilities, despite such pay being a required element of TIF grants.

Endnotes

- 1 Allan Odden and Carolyn Kelley, *Paying Teachers for What They Know and Do: New and Smarter Compensation Strategies to Improve Schools*, Second Edition (Thousand Oaks, CA: Corwin Press, 2002).
- 2 Joan Baratz-Snowden, "The Future of Compensation Reform: Déjà Vu or Something New?" (Washington: Center for American Progress, 2007).
- 3 Lynn Olson, "Teacher-Pay Experiments Mounting Amid Debate," *Educa-tion Week*, October 3, 2007, available at <http://www.edweek.org/ew/articles/2007/10/03/06pay.h27.html?tmp=1553238413>.
- 4 Baratz-Snowden, "The Future of Compensation Reform."
- 5 Robin Chait and Reagan Miller, "Paying Teachers for Results: A Summary of Research to Inform the Design of Pay-for-Performance Programs for High-Poverty Schools" (Washington: Center for American Progress, 2008).
- 6 Herbert G. Heneman, Anthony Milanowski, and Steven Kimball, "Teacher Performance Pay: Synthesis of Plans, Research, and Guidelines for Practice" (Madison, WI: University of Wisconsin-Madison, 2007).
- 7 Allan Odden, "New Teacher Pay Structures: The Compensation Side of the Strategic Management of Human Capital" (Madison, WI: University of Wisconsin-Madison, 2008).
- 8 Ibid.
- 9 National Center for Analysis of Longitudinal Data in Education Research, "Plain Talk with Dan Goldhaber," available at <http://www.caldercenter.org/news/ptD-Goldhaber.cfm>.
- 10 Robin Chait, "Current State Policies that Reform Teacher Pay: An Examination of Pay-for-Performance Programs in Eight States" (Washington: Center for American Progress, 2007).
- 11 Innovation Districts for School Improvement Act, S. 114, 110 Cong. 1 sess.
- 12 Teacher Excellence for All Children Act of 2007, S. 1339, 110 Cong. 1 sess.
- 13 Congressman George Miller and Congressman Buck McKeon, "Summary of Miller-McKeon Discussion Draft, Title II—Teacher Excellence for All Children (TEACH)."
- 14 Herbert G. Heneman and Anthony Milanowski, "Assessing Human Resource Alignment: The Foundation for Building Total Teacher Quality Improvement" (Madison, WI: University of Wisconsin-Madison, 2007).
- 15 Ibid.
- 16 Ibid.
- 17 Charlotte Danielson, *Enhancing Professional Practice: A Framework for Teaching*, 2nd Edition (Alexandria, VA: Association for Supervision and Curriculum Development, 2007). Robert C. Pianta, Karen M. La Paro, and Bridget K. Hamre, *Classroom Assessment Scoring System Manual, K-3* (Baltimore: Paul H. Brookes Publishing Company, 2008). The CLASS manual for secondary grades must be ordered directly through the program's website at <http://www.classobservation.com/>. For a summary and comparison of those competency frameworks along with five others, see Anthony T. Milanowski and Herbert G. Heneman, "Review of Teaching Performance Assessments for Use in Human Capital Management" (Madison, WI: University of Wisconsin-Madison, 2009).
- 18 Odden, "New Teacher Pay Structures."
- 19 Heneman and Milanowski, "Assessing Human Resource Alignment."
- 20 Importantly, in the private sector "evaluation" includes not just the act of assessing performance but also using that information for "performance management"—providing feedback and coaching to employees, rewarding or promoting strong performers, helping weak performers improve, and reassigning or dismissing persistently poor ones.
- 21 Author's estimate based on figures in U.S. Department of Education, *Findings From the 2007-08 Survey on the Use of Funds Under Title II, Part A* (2008).
- 22 Allan Odden and James A. Kelly, "Strategic Management of Human Capital in Public Education" (Madison, WI: University of Wisconsin-Madison, 2008).
- 23 Douglass N. Harris and Tim R. Sass, "Teacher Training, Teacher Quality, and Student Achievement" (Washington: National Center for Analysis of Longitudinal Data in Education Research, 2007).
- 24 Linda Darling-Hammond and others, "Professional Learning in the Learning Profession: A Status Report on Teacher Development in the United States and Abroad" (Dallas: National Staff Development Council, 2009).
- 25 U.S. Department of Education, "Findings from the 2007-08 Survey on the Use of Funds Under Title II, Part A."
- 26 Karen Hawley-Miles, "Turning Tough Times into Opportunity: Finding the Resources We Need to Build Teacher Quality and Close The Achievement Gap" (Education Resource Strategies, 2008).
- 27 Darling-Hammond and others, "Professional Learning in the Learning Profession."
- 28 Kwang Suk Yoon and others, *Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement* (Washington: U.S. Department of Education, 2007).
- 29 Ibid.
- 30 The first was Michael S. Garet and others, "What Makes Professional Development Effective? Results from a National Sample of Teachers," *American Educational Research Journal* 38 (4) (2001): 915-945. The second was Laura M. Desimone and others, "Effects of Professional Development on Teachers' Instruction: Results from a Three-Year Longitudinal Study," *Educational Evaluation and Policy Analysis* 24 (2) (2002): 81-112.
- 31 Garet and others, "What Makes Professional Development Effective?"
- 32 Karen Hawley-Miles and Stephen Frank, *The Strategic School: Making the Most of People, Time, and Money* (Thousand Oaks, CA: Corwin Press, 2008).
- 33 Andrew Wayne and others, "Experimenting with Teacher Professional Development: Motives and Methods," *Educational Researcher* 37 (8): 469-479.
- 34 Author's analysis of the U.S. Schools and Staffing Survey 2003-4, Data Analysis System, available at <http://nces.ed.gov/das/>.
- 35 Aurora D'Amico, unpublished data from the 2003-4 U.S. Schools and Staffing Survey.
- 36 Elizabeth A. City and others, *Instructional Rounds in Education: A Network Approach to Improving Teaching and Learning* (Cambridge, MA: Harvard Education Press, 2009).
- 37 Education Resource Strategies, "Teaching Quality: The First Priority" (2009).
- 38 City and others, *Instructional Rounds in Education*.
- 39 Julie Kowal and Lucy Steiner, "Instructional Coaching" (Washington: The Center for Comprehensive School Reform and Improvement, 2007).

- 40 James E. Taylor, "Instructional Coaching: The State of the Art." In Melinda M. Mangin and Sara Ray Stoelinga, eds., *Effective Teacher Leadership: Using Research to Inform and Reform* (New York: Teachers College Press, 2008).
- 41 North Dakota Department of Education website, "Title I Programs: Instructional Coaches," available at <http://www.dpi.state.nd.us/title1/coaches.shtml>.
- 42 Melinda M. Mangin and Sara Ray Stoelinga, "Teacher Leadership: What It Is and Why It Matters." In Melinda M. Mangin and Sara Ray Stoelinga, eds., *Effective Teacher Leadership: Using Research to Inform and Reform* (New York: Teachers College Press, 2008).
- 43 Mangin and Stoelinga, "Teacher Leadership."
- 44 Brian Lord, Kate Cress, and Barbara Miller, "Teacher Leadership in Support of Large-Scale Mathematics and Science Education Reform." In Melinda M. Mangin and Sara Ray Stoelinga, eds., *Effective Teacher Leadership: Using Research to Inform and Reform* (New York: Teachers College Press, 2008).
- 45 Ibid.
- 46 Sarah Murray, Xin Ma, and Joan Mazure, "Effects of Peer Coaching on Teachers' Collaborative Interactions and Students' Mathematics Achievement," *Journal of Educational Research* 102 (3) (2009): 203-212.
- 47 Daniel T. Willingham, *Why Don't Students Like School? A Cognitive Scientist Answers Questions About How the Mind Works and What It Means for the Classroom* (New York: Jossey-Bass, 2009).
- 48 K. Anders Ericsson, "The Influence of Experience and Deliberate Practice on the Development of Superior Expert Performance." In K. Anders Anderson and others, eds., *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge, United Kingdom: Cambridge University Press, 2006).
- 49 Willingham, *Why Don't Students Like School?*
- 50 Ericsson, "The Influence of Experience and Deliberate Practice on the Development of Superior Expert Performance."
- 51 Ibid.
- 52 Darling-Hammond and others, "Professional Learning in the Learning Profession."
- 53 Cheryl L. Kirkpatrick, "To Invest, Coast or Idle: Second-Stage Teachers Enact their Job Engagement" (Paper presented at the American Educational Research Association Annual Conference, Chicago, IL, April 10, 2007).
- 54 K. Anders Ericsson, ed., *Development of Professional Expertise: Toward Measurement of Expert Performance and Design of Optimal Learning Environments* (Cambridge, UK: Cambridge University Press, forthcoming).
- 55 Kirkpatrick, "To Invest, Coast or Idle."
- 56 Thomas Toch and Robert Rothman, "Rush to Judgment: Teacher Evaluation and Public Education" (Washington: Education Sector, 2008).
- 57 Jane Hannaway and Andrew J. Rotherham, "Collective Bargaining in Education and Pay for Performance" (Nashville, TN: National Center on Performance Incentives, 2008).
- 58 National Council on Teacher Quality, "National Council on Teacher Quality State Teacher Policy Yearbook: What States Can Do To Retain Effective New Teachers" (2009).
- 59 Hannaway and Rotherham, "Collective Bargaining in Education and Pay for Performance."
- 60 National Council on Teacher Quality, "National Council on Teacher Quality State Teacher Policy Yearbook."
- 61 Quoted in Toch and Rothman, "Rush to Judgment."
- 62 Daniel Weisberg and others, "The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness" (New York: The New Teacher Project, 2009).
- 63 Ibid.
- 64 Ibid.
- 65 Megan W. Stuhlmiller and Robert C. Pianta, "Profiles of Educational Quality in First Grade," *Elementary School Journal* 109 (4) (2009): 323-342.
- 66 Daniel Weisberg and others, "The Widget Effect."
- 67 Ibid.
- 68 Ibid.
- 69 Ann Duffett and others, "Waiting to Be Won Over: Teachers Speak on the Profession, Unions, and Reform" (Washington: Education Sector, 2008).
- 70 Daniel Weisberg and others, "The Widget Effect."
- 71 Susan P. Choy and others, "Teacher Professional Development in 1999-2000" (Washington: National Center for Education Statistics).
- 72 Letter from U.S. Secretary of Education Arne Duncan to Governors, April 1, 2009.
- 73 U.S. Department of Education, *American Recovery and Reinvestment Act of 2009: Using ARRA Funds to Drive School Reform and Improvement* (April 24, 2009).
- 74 Anthony T. Milanowski and Herbert G. Heneman, "Assessment of Teacher Reactions to a Standards-Based Teacher Evaluation System: A Pilot Study," *Journal of Personnel Evaluation in Education* 15 (3) (2001): 193-212.
- 75 Steven M. Kimball, "Analysis of Feedback, Enabling Conditions, and Fairness Perceptions of Teachers in Three School Districts with New Standards-Based Evaluation Systems," *Journal of Personnel Evaluation in Education* 16 (4) (2002): 241-268.
- 76 Richard Halverson, Carolyn Kelley, and Steven Kimball, "Implementing Teacher Evaluation Systems: How Principals Make Sense of Complex Artifacts to Shape Local Instructional Practice." In Wayne K. Hoy and Cecil G. Miskel, eds., *Educational Administration, Policy and Reform: Research and Measurement* (Greenwich, CT: Information Age Publishing, 2004).
- 77 Ibid.
- 78 Regis Anne Shields and Karen Hawley Miles, "Strategic Designs: Lessons from Leading Edge Small Urban High Schools" (Watertown, MA: Education Resource Strategies, 2008).
- 79 Ibid.
- 80 Toch and Rothman, "Rush to Judgment."
- 81 Ibid.
- 82 Harvard University Project on the Next Generation of Teachers, "A User's Guide to Peer Assistance and Review," available at <http://www.gse.harvard.edu/~ngt/par/practice/toledo.html>.
- 83 Ibid.
- 84 Allan Odden, interview with author, April 3, 2009.
- 85 National Institute for Excellence in Teaching, available at <http://www.talent-edteachers.org/action/action.taf?page=where>
- 86 Monica Knauer, interview with author, Los Angeles, CA, April 17, 2009.
- 87 John Schacter and Yeow Meng Thum, "Paying for High- and Low-Quality Teaching," *Economics of Education Review* 23 (2004): 411-430.
- 88 Knauer, interview with author.
- 89 Eric Matheson, panel presentation on "Implementing an Evaluation System That Is Fair and Transparent," Ninth Annual Teacher Advancement Program Conference, Los Angeles, Saturday April 18.
- 90 Barron, interview with author, April 6, 2009.
- 91 Ericsson, "The Influence of Experience and Deliberate Practice on the Development of Superior Expert Performance."
- 92 Alma Velez, interview with author, Los Angeles, CA, April 17, 2009.
- 93 Lynn Kuykendall, interview with author, April 2, 2009.
- 94 For teachers not assigned to grades or subjects where classroom gains can be calculated, TAP schools typically increase the weight of the schoolwide gain factor, though other variations exist.
- 95 Melodie Barron, interview with author.

- 96 Cassandra Slayton, interview with author, Los Angeles, CA, April 18, 2009.
- 97 Kuykendall, interview with author.
- 98 Velez, interview with author.
- 99 Vicki A. Condalary, interview with author, Los Angeles, CA, April 18, 2009.
- 100 Kuykendall, interview with author.
- 101 Heneman, Milanowski, and Kimball, "Teacher Performance Pay."
- 102 Calculations by the author.
- 103 Herbert G. Heneman, "Strategic Management of Human Capital in Minnesota: Quality Compensation for Teachers (Q-Comp)" (Madison, WI: University of Wisconsin-Madison, 2008).
- 104 Susan Moore Johnson, Jill Harrison Berg, and Morgan L. Donaldson, "Who Stays In Teaching and Why: A Review of the Literature on Teacher Retention" (Washington: NRTA Educator Support Network, 2005).
- 105 Charles Clotfelter and others, "Would Higher Salaries Keep Teachers in High-Poverty Schools? Evidence from a Policy Intervention in North Carolina," *Journal of Public Economics* 92 (2008): 1352-1370.
- 106 Jessica Levin and Meredith Quinn, "Missed Opportunities: How We Keep High-Quality Teachers Out of Urban Classrooms" (New York: The New Teacher Project, 2003).
- 107 Susan Moore Johnson and Morgan L. Donaldson, "Sustaining New Teachers Through Professional Growth." In Susan Moore Johnson, ed., *Finders and Keepers: Helping New Teachers Survive and Thrive in Our Schools* (San Francisco: Jossey-Bass, 2004).
- 108 Public Agenda, "An Assessment of Survey Data on Attitudes about Teaching Including the Views of Parents, Administrators, Teachers and the General Public" (2003).
- 109 Richard Ingersoll and David Perda, "The Mathematics and Science Teacher Shortage: Fact and Myth" (Philadelphia: Center for Policy Research in Education, 2009).
- 110 Vincent M. Mallozzi, "Career Ladders for Teachers on the Rise," *The New York Times*, April 10, 1988, available at <http://www.nytimes.com/1988/04/10/education/blackboard-notes-career-ladders-for-teachers-on-the-rise.html>.
- 111 Odden and Kelley, Paying Teachers for What They Know and Do.
- 112 Missouri Department of Elementary and Secondary Education, "Criteria for Local District Career Ladder Plans," available at <http://dese.mo.gov/divteachqual/careerladder/appenda.html> (last accessed July 1, 2009).
- 113 Missouri Department of Elementary and Secondary Education, "Missouri's Career Ladder Program: Annual Report 2007-2008" (2008).
- 114 S.B. 1521 - 1990 (Chapter 319) Section 15-918-02 Career ladder program; requirements; optional incentive component.
- 115 Author's analysis of individual "career ladder handbooks" published by districts.
- 116 Amphitheater Public Schools, Career Ladder Program, available at <http://www.amphi.com/departments/careerladder/home.html>
- 117 Author's analysis of individual "career ladder handbooks" published by districts.
- 118 Jane Dowling and others, "The Effects of the Career Ladder Program on Student Achievement" (Phoenix, AZ: Sheila Murphy Associates, 2007).
- 119 Kevin Booker and Steven Glazerman, "Missouri's Teacher Career Ladder Program" (Nashville, TN: National Center on Performance Incentives, 2008).
- 120 Minnesota Department of Education, "Application for Q Comp Aid."
- 121 Beth Driscoll, interview with author, May 11, 2009.
- 122 Kristan Van Hook, interview with author, June 8, 2009.
- 123 Craig D. Jerald, "Benchmarking for Success: Ensuring U.S. Students Receive a World-Class Education" (Washington, DC: National Governors Association, Council of Chief State School Officers, and Achieve, Inc., 2008).
- 124 Craig D. Jerald, "Keeping Kids in School: What Research Says about Preventing Dropouts" (Alexandria, VA: National School Boards Association, Center for Public Education).
- 125 Morgan L. Donaldson and others, "Angling for Access, Bartering for Change: How Second-Stage Teachers Experience Differentiated Roles in Schools," *Teachers College Record* 110 (5) (2008): 1088-1114.
- 126 Ibid.
- 127 Richard F. Elmore, "Bridging the Gap Between Standards and Achievement" (Washington: Albert Shanker Institute, 2002).
- 128 Beverly Showers, "Teachers Coaching Teachers," *Educational Leadership* 42 (7) (1985): 43-48.
- 129 Jennifer Goldstein, "Debunking the Fear of Peer Review: Combining Supervision and Evaluation and Living to Tell About It," *Journal of Personnel Evaluation in Education* 18 (2005): 235-252.
- 130 Lord, Cress, and Miller, "Teacher Leadership in Support of Large-Scale Mathematics and Science Education Reform."
- 131 Eric M. Camburn, Steven M. Kimball, and Rebecca Lowenhaupt, "Coaching to Scale with Teacher Leadership: Lessons Learned from a Districtwide Literacy Coach Initiative." In Melinda M. Mangin and Sara Ray Stoelinga, eds., *Effective Teacher Leadership: Using Research to Inform and Reform* (New York: Teachers College Press, 2008).
- 132 Sara Ray Stoeling and Melinda M. Mangin, "Drawing Conclusions about Instructional Teacher Leadership." In Melinda M. Mangin and Sara Ray Stoelinga, eds., *Effective Teacher Leadership: Using Research to Inform and Reform* (New York: Teachers College Press, 2008).
- 133 Jennifer Goldstein, "Easy to Dance To: Solving the Problems of Teacher Evaluation with Peer Assistance and Review," *American Journal of Education* 113 (2007): 479-508.
- 134 Susan Sclafani, "Rethinking Human Capital in Education: Singapore as a Model for Teacher Development" (Washington: The Aspen Institute, 2008).
- 135 Ibid.
- 136 Michael Barber and Mona Mourshed, "How the World's Best-Performing School Systems Come Out on Top" (London: McKinsey and Company, 2007).
- 137 Sclafani, "Rethinking Human Capital in Education."
- 138 Allan Odden, interview with author.
- 139 Marcus Winters and others, "The Effect of Performance Pay in Little Rock, Arkansas on Student Achievement" (Nashville, TN: National Center on Performance Incentives, 2008).
- 140 U.S. Department of Education, "Teacher Incentive Fund" available at <http://www.ed.gov/programs/teacherincentive/faq.html> (last accessed July 1, 2009).
- 141 U.S. Department of Education, "Notice Inviting Applications for New Awards for Fiscal Year (FY) 2006" available at <http://www.ed.gov/legislation/fedRegister/announcements/2006-4/111406a.html> (last accessed July 1, 2009).
- 142 U.S. Department of Education, "Fiscal Year 2010 Budget Summary" (May 7, 2009), available at <http://www.ed.gov/about/overview/budget/budget10/summary/edlite-section3a.html#tif>.
- 143 The Center for Educator Compensation Reform, <http://www.cecr.ed.gov/>.

About the author

Craig Jerald is president of Break the Curve Consulting, which provides expertise to leaders and policymakers on issues related to education policy, communications, research, and practice. Since founding Break the Curve in 2004, Craig has worked with organizations such as the Center for American Progress, the Bill and Melinda Gates Foundation, the National Governors Association, the National School Boards Association, the Education Trust, Achieve Inc., Education Sector, the Learning First Alliance, ACT, and the Alliance for Excellent Education. Prior to founding Break the Curve, Craig served as a principal partner at the Education Trust, where he worked on issues ranging from school accountability to teacher quality and school finance equity. As a senior editor at *Education Week* from 1996 to 2000, Craig founded and managed the organization's research division and helped create and direct *Education Week*'s special annual reports series, *Quality Counts* and *Technology Counts*. Craig also has worked at the U.S. Department of Education, and he began his career as a Teach for America recruit and middle school teacher in California's Long Beach Unified School District.

Acknowledgements

The Center for American Progress thanks the Joyce Foundation for generously providing support for this paper. Based in Chicago, the foundation invests approximately \$8 million annually in work to improve public education in the Midwest, especially by improving the quality of teachers in low-performing schools.

The author would like to thank a number of individuals who took time to offer their expertise on this issue. Lowell Milken provided an invaluable founder's perspective on the history of the Teacher Advancement Program and lessons learned over the past decade. Kristan Van Hook, senior vice president of the National Institute for Excellence in Teaching, answered innumerable questions about how the TAP model works and offered trenchant perspectives on many teacher workforce issues, along with introductions to NIET colleagues who provided additional expertise, including Gary Stark, Melodie Barron, and Stephanie Sydia.

The author especially wishes to thank the TAP Master Teachers who agreed to share their personal experiences and insights about these issues, including Sherry Cashion, Monica Knauer, Lynn Kuykendall, Cassandra Slayton, and Alma Velez. Two Executive Master Teachers with the Louisiana Teacher Advancement Program, Vicki Condalary and Susan Couch, also provided useful information and perspectives on TAP. Beth Driscoll provided information about Arizona's career ladder districts and helped me understand the major challenges and key questions policymakers confront when designing programs to pay teachers for assuming additional roles and responsibilities. Aurora D'Amico provided invaluable assistance in accessing data from the federal Schools and Staffing Survey. Allan Odden provided a wealth of information on emerging strategies for managing human resources in education and critical insights about how those strategies can advance the broader goal of school improvement. Finally, the author offers special thanks to Cindy Brown and Robin Chait of the Center for American Progress for their infinite patience and wise counsel as he explored the many complex issues related to this topic.

The Center for American Progress is a nonpartisan research and educational institute dedicated to promoting a strong, just and free America that ensures opportunity for all. We believe that Americans are bound together by a common commitment to these values and we aspire to ensure that our national policies reflect these values. We work to find progressive and pragmatic solutions to significant domestic and international problems and develop policy proposals that foster a government that is "of the people, by the people, and for the people."

