Teacher pay-for-performance: A framework for program design

FINAL DRAFT

Forthcoming paper in the

Economic Policy Institute Series on Alternative Teacher Compensation Systems

John P. Papay and Susan Moore Johnson

Harvard University Graduate School of Education
Project on the Next Generation of Teachers

December 2008
Teacher pay-for-performance: A framework for program design

Introduction

While the single salary scale, which rewards teachers based on classroom experience and education level, remains ubiquitous in America’s public schools, many states, districts, and schools are now experimenting with alternative compensation strategies. Pay reforms are not new in education: the country has experienced at least three past waves of enthusiasm for merit pay, the most recent following the release of *A Nation at Risk* in 1983. Today, experiments are enjoying a strong resurgence, driven in part by standards-based reform and educational research showing that teachers contribute substantially to student learning.¹ Policy makers and politicians from across the political spectrum have endorsed or championed differentiated pay approaches (e.g., American Federation of Teachers 2008; Obama 2008; Fordham Foundation 2008) and districts and states have spent hundreds of millions of dollars to reform teacher compensation. Beginning in 2006, the federal government joined in, allocating nearly $100 million for the Teacher Incentive Fund (TIF), which “supports efforts to develop and implement performance-based teacher and principal compensation systems in high-need schools” (U.S. Department of Education 2008).

These compensation reforms have various names—differentiated pay, pay-for-performance, professional compensation, merit pay, and performance-based pay—but all supplement the traditional salary scale with bonuses for teachers that range from a few hundred to several thousands of dollars. They use vastly different criteria, however, in determining how to allocate the awards. In this review, we focus primarily on programs—which we term “pay-for-performance”—offering bonuses for effective teaching performance, evidenced by formal

¹ The literature on the importance of teachers and the variation in teacher effectiveness has grown rapidly in recent years (see for example Gordon, Kane, and Staiger (2006); Rivkin, Hanushek, and Kain (2005); Rockoff (2004); Aaronson, Barrow, and Sander (2003)).
evaluations or student test scores. These programs differentiate pay for teachers on the basis of their success in the classroom, rather than for their knowledge and skills, the role they have in the school, or the teaching assignment they take.

Even among programs that include “pay-for-performance,” there remains substantial variation in both the design and implementation of state and local initiatives. We present a framework that details the key ways pay-for-performance programs differ. This framework centers on three of the most important design questions that districts must answer: (1) what measures—standardized tests, professional evaluation, or several indicators—to use in assessing performance; (2) whether to identify top performers with relative rankings or standards-based comparisons; and (3) whether to provide awards at the individual or group level. Each of these decisions presents important challenges and produces a distinct set of incentives for teachers that differ in their focus, strength, and effects.

Although program design plays an important role in a program’s success, a variety of other factors, including the implementation and local context, matter a great deal. Examining pay-for-performance components in isolation can provide only a partial view of how these programs work in local districts. Therefore, we extend this theoretical discussion with descriptions of pay-for-performance programs in four large, urban districts—Houston, Texas, Minneapolis, Minnesota, Charlotte-Mecklenburg, North Carolina, and Hillsborough County, Florida—all of which provide substantial awards for teacher performance. These examples are useful because they all supplement state-operated pay-for-performance plans with local programs. These districts—along with others that have well-publicized initiatives, such as Denver, Colorado, Douglas County, Colorado, and Toledo, Ohio—offer several pay-for-performance components, providing teachers multiple opportunities to earn bonuses.
We begin our discussion with an overview of teacher compensation and a description of compensation reforms currently in use across the country. Next, we present our framework, illustrating the key dimensions that distinguish different pay-for-performance plans. We then illustrate this framework with examples of program elements from different districts. We conclude with a discussion of several key themes that emerge in our exploration of these programs.

1.1 Teacher compensation: An overview

The single salary scale is among the most prevalent structures in public education. Starting in 1921 in Des Moines and Denver, it arose in response to charges of inequity and administrative abuse (Odden and Kelley 2002). It flourished in mid-century and, when collective bargaining first began for public school teachers in the 1960s, unions and school boards typically agreed to preserve this structure. In 2003-04, over 93% of school districts paid teachers according to such a scale (Strizek et al. 2003-04). In most cases, teachers increase their pay over time by moving up “steps” for experience in the district (and/or the profession) and across “lanes” by completing coursework and/or advanced degrees. All teachers with the same combination of experience and education receive the same salary without regard to their classroom performance. This scale provides strong incentives for teachers to remain in the district and to accrue professional credits and degrees, essentially rewarding loyalty with stable, steadily increasing compensation.

Merit pay, in fact, has an even longer history in American education. Following Frederick Taylor’s “scientific management” movement in the late 19th and early 20th centuries, many school district officials experimented with granting salary increases using administrators’
assessments of individual teachers’ merit or effectiveness (Evenden 1918; Young 1933). Rife with administrative abuse, these plans quickly disappeared and were replaced by the single salary scale (Johnson 1984). The challenge to American education provoked by Sputnik produced another short-lived series of merit pay proposals in the 1960s. Subsequently, the release of *A Nation at Risk* in 1983, which sparked the standards-based reform movement, turned attention to defining what students should learn in schools and to assessing this learning. Along with education researchers’ conclusions that easily measurable educational inputs had little effect on student achievement (e.g., Hanushek 1986), this report helped spark the fleeting “merit pay” movement in the 1980s. In each of these periods, merit pay rose quickly to prominence and then faded rapidly, only to have reformers decades later view the approach as a novel reform (Johnson 1986).

The failings of these past merit pay efforts have been well documented (Murnane and Cohen 1986; Johnson 1986; Odden and Kelley 2002; Podgursky and Springer 2007). These programs often included small awards, lacked transparency, and based decisions on cursory evaluations. In many cases, teachers did not know what they needed to do in order to achieve a performance bonus (Murnane and Cohen 1986). As a result, these plans did not provide sufficient incentives to change teachers’ practice and often provoked resentment and opposition from educators. Furthermore, a lack of stable funding often doomed these programs to failure because they became easy targets during budget cuts.

The current iteration of differentiated pay proposals faces similar challenges. Securing stable funding remains an important, and demanding, pre-requisite. Although states may enjoy greater access to financial resources than local districts, many state-funded programs have suffered from one-size-fits-all implementation. Different districts face different educational
problems, suggesting that the incentives that are effective in one location may not work in another. In addition, the political, cultural, and organizational realities of local schools mean that some plans may be viable in certain locations but not others.

**The Landscape of Compensation Reforms**

Recent calls for compensation reform derive largely from the changing educational context over the past decade. Increasing regulation and accountability, from sources like the No Child Left Behind Act of 2001, and competition, from charter schools and school choice, have increased pressures for districts to improve student achievement. Educational research has also confirmed what parents have long known: teachers vary substantially in their contributions to student achievement (e.g., Gordon, Kane, and Staiger 2006; Rivkin, Hanushek, and Kain 2005; Rockoff 2004; Aaronson, Barrow, and Sander 2003). At the same time, the educational labor force is undergoing rapid transition as Baby Boom teachers who spent their entire career in the classroom retire. Although veteran teachers remember the failed attempts at merit pay and tend to respond with skepticism, newer teachers often express greater degrees of support for differentiating salaries on the basis of performance (Jacob and Springer, 2008; Johnson et al., 2007; Goldhaber, DeArmond, and DeBurgomaster, 2007).

Thus, there are increasing efforts to reform the single salary scale. In 2003-04, about 8% of public schools reported using a pay incentive to “reward excellence in teaching” (Strizek et al. 2006), up 25% since 1999-2000 (Podgursky and Springer 2007). Initiatives in pay reform are increasingly widespread. Many states have developed programs or pilots to implement some differentiated pay reforms (National Center for Performance Incentives 2008).² Districts are also designing incentive programs on their own; recently, New York City, the country’s largest school district, announced a performance-based pay program for its 80,000 teachers (Gootman

---

² Texas, Minnesota, Florida, Alaska, Iowa, North Carolina, Missouri, and Kentucky have all introduced such plans.
Importantly, nearly all of these proposals build on the single salary scale by offering bonuses or salary increments that supplement a standardized base salary.\(^3\)

These reform proposals come in many varieties and names. Many scholars have written about the range of teacher pay initiatives (e.g., Podgursky and Springer 2007; Odden and Kelley 2002). Here, we divide approaches to reforming the single salary scale into four main categories:

1. **Knowledge and skills**: pay for undertaking professional development or acquiring skill-based credentials
2. **Roles**: pay for assuming special roles and responsibilities
3. **Market factors**: pay for teaching in hard-to-staff subjects or schools
4. **Performance**: pay for effective instructional practice and student achievement

Each of these efforts is designed to improve the quality and performance of teachers by using financial incentives that reward more than loyalty and educational attainment. They encourage teachers to undertake new learning relevant to instructional practice, to take on different roles or assignments, and to implement more effective practices or exert greater effort in their classrooms. By altering the incentive structures in teacher compensation, districts strive to change teacher behaviors and improve student performance.

**Knowledge and skills-based policies** reward teachers for improving their pedagogical content knowledge and instructional skills. Unlike lane advances on the traditional salary scale, these reforms typically offer salary increases for employees who pursue professional development that ties directly to classroom instruction or school leadership (see Odden and Kelley 2002). Many states and districts now offer bonuses to teachers certified by the National

---

\(^3\) Only Denver, which implemented its ProComp program in 2005, has substantially revised the structure of teacher compensation.
Board for Professional Teaching Standards (NBPTS) or to those demonstrating pedagogical knowledge and skill through a comprehensive evaluation process.

Building on the existing structure of providing supplemental contracts for athletic coaches and activity advisors, role-based policies differentiate pay for teachers by their responsibilities. Teachers typically receive a stipend and release time from classroom activities to serve as instructional coaches, consulting teachers, or lead teachers within a school. These role-based reforms often coincide with career ladders, which provide a staged career through which teachers proceed, for example, from novice to career teacher to lead teacher.

Market-based compensation represents a further step away from the single salary scale because it rewards certain classroom teachers over others based on their subject or school. Many districts—particularly urban school systems—struggle to fill positions in certain subject areas, such as special education, math, and science. Similarly, many districts have particular schools that, because of working conditions or ineffective leadership, experience dramatic turnover and consistent vacancies. In response to these challenges, districts have begun providing bonuses to counter the labor supply factors that lead to these vacancies in hard-to-staff subjects and/or schools.

Finally, states and local districts are rewarding successful teachers based on their instructional practice, as reflected in student test scores or professional evaluations. These plans come closest to earlier “merit pay” initiatives. In this paper, we concentrate on these programs because they remain the most challenging and least understood of compensation reforms. In fact, they are among the most contentious policy initiatives in public education today and represent an area of great current interest for policy makers. This focus, however, is somewhat artificial since

4 For example, 35% of urban schools had trouble filling math vacancies in 2003-04, compared to just 8% in English and 5% in social studies (Strizek et al. 2004).
many successful plans include elements from several different categories at the same time, sometimes integrating them into a broader strategy for human capital management. Moreover, the lines between these approaches often are blurred. For example, career ladders award pay differentially based on roles and responsibilities as well as knowledge and skills, but decisions about promotion up the ladder typically involve assessments of a teacher’s performance through professional evaluation. We return to these issues at the end of the paper.

1.2 Pay-for-performance: a framework of approaches

Theoretical background of pay-for-performance

Advocates of pay-for-performance plans rely on extrinsic incentives to change individuals’ behavior. As such, these plans implicitly derive from a theory of motivation called expectancy theory. First articulated in the 1960s, expectancy theory argues that the motivational effect of an incentive depends on the value individuals place on the reward and their beliefs about the likelihood of attaining it (see Vroom 1964; Lawler 1983). Clearly, larger and more valuable rewards serve as more effective motivators—for a plan to have its desired effects, bonuses must be sufficiently large to influence behavior.5 Determining the appropriate number of awards remains a related, but more difficult, challenge. Here, teachers must reasonably expect that they will achieve the reward if they put forth additional effort, but the awards must not be so widespread that most teachers earn them with only minimal increases in effectiveness.

Within a fixed budget, then, districts inevitably face tradeoffs: a larger award necessarily means that fewer teachers will receive the bonus, and districts must strike a balance between the award being too scarce and too small. These decisions are key to the success of any performance

---

5 Evidence from other fields suggests that small incentives may in fact produce resentment and decrease effort (Gneezy and Rustichini 2000).
incentives in education. Pay-for-performance programs deliberately manipulate extrinsic incentives to produce changes in behavior in a field where the work is driven, in large part, by intrinsic motivation (Lortie, 1975). Awards may backfire if they reduce teachers’ intrinsic motivation, yet are not sufficient extrinsic motivators. This concern is particularly acute if the awards are not combined with good working conditions.

Advocates of pay-for-performance argue that the single salary scale decouples the largest school expenditure, teachers’ salaries, from public education’s most important outcome, student learning. They assert that introducing more output-based compensation will improve teacher quality through two mechanisms: by motivating current teachers and by changing the composition of the teaching workforce. First, monetary rewards for student performance will motivate existing teachers to improve their instructional practice and exert additional effort. Here, not only could pay-for-performance motivate teachers to work harder in the classroom, but it could also encourage them to pursue additional professional development and direct these activities more effectively to improve student outcomes. Second, the overall quality of the teacher workforce will increase as high-quality teachers, who expect to get awards, enter and stay in the profession while low-quality teachers, who do not expect to receive the awards, exit. Thus, pay-for-performance is designed to attract and retain high performing teachers. Advocates argue that this selection mechanism is particularly important today as cognitive demands on teachers have risen, while the overall quality of the entering teaching force has not (Corcoran, Evans, and Schwab 2004; Hoxby and Leigh 2004).

Opponents of pay-for-performance argue that, unlike manufacturing, schooling is a highly personalized craft that requires ongoing adaptation for different groups of students and settings, thus limiting the use of outcomes-based compensation. First, the production process in
education is quite complex and isolating an individual teacher’s contribution to student learning remains a difficult – some say impossible – empirical task. Teachers are but one of many contributors to student learning. Beyond the classroom, a wide variety of in-school and out-of-school factors affect how much students learn in a single year (see Rothstein, 2004). Education is also, in many ways, a group production process where teachers of multiple grades and subjects contribute to a child’s education. Those who oppose pay-for-performance also argue that such initiatives would promote unproductive competition among teachers, who need to collaborate, and generate an atmosphere that is hostile to learning. Furthermore, the ultimate instructional outcome that should be assessed remains unclear: achievement tests cannot capture the wide range of things that schools do and the public values. Finally, even if an individual teacher’s contribution to student achievement could be isolated, there is no way to determine the monetary value of a specific increase in achievement (Neal 2007).

Opponents also argue that the primary theories of action underlying pay-for-performance ignore the realities of schooling and may not be the most effective policy solutions. Asserting that pay-for-performance works by increasing teachers’ motivation implicitly suggests that teachers could improve student learning simply by working harder. If, instead, the real challenge facing schools is a lack of instructional capacity (Elmore, 2004), this policy will have little effect unless it is coupled with intensive and effective professional development. While pay-for-performance could attract a stronger pool of applicants with better basic communication skills or subject-matter knowledge, other strategies, such as professional development, would be needed to enhance the capacity of the current teaching force.
A framework of pay-for-performance programs

Implementing pay-for-performance programs requires districts to make important decisions about program design and implementation. Resolving implementation challenges – e.g., securing stable funding, obtaining buy-in from relevant stakeholders, and deciding whether to allow voluntary participation in the program – is essential for the program to succeed. However, the design of these programs is also crucial. Districts must determine the size of awards, the type of compensation, whether to provide one-time bonuses or permanent increases, the source of funds, and a whole host of other matters. Here, we focus on the three most important decisions:

- **How to measure performance**: standardized tests, professional evaluation, or mixed measures?
- **How to identify top performers**: relative rankings or standards-based comparisons?
- **At what level to provide rewards**: individual or group?

Using these three dimensions, we present a framework of pay-for-performance programs and twelve main “types” of pay-for-performance programs, as seen in Figure A. The literature on performance-based pay identifies many strengths and weaknesses inherent in each choice; no one type is without challenges. In this section, we address the theoretical implications underlying each of these dimensions.

(1) **How to measure performance: Tests, evaluations, or mixed measures**

In order to reward teachers for their performance, districts must first define what they mean by “performance” and how to measure it. Here, districts have several main options: they can use student achievement test scores, professional evaluations, or multiple measures. Today, most districts with pay-for-performance plans reward teachers for standardized test scores through some sort of student growth measure. However, several districts use professional
evaluation, typically conducted by district administrators, or they combine evaluations and test scores into a composite measure. Each of these options comes with important advantages and disadvantages.

In the context of standards-based reform, student test scores have become the most important outcome by which schools are held accountable. Thus, offering teachers incentives for raising student test scores may be an effective way to meet accountability pressures and to improve performance on fundamental educational outcomes. Furthermore, student test scores are arguably objective measures of teachers’ performance, void of personal biases. Thus, they maintain some of the virtues of the single salary scale’s elimination of favoritism and administrative abuse.

For districts planning to use test scores, most analysts recommend focusing attention on student growth, or “value-added”, rather than achievement levels, because providing incentives for teachers whose students achieve high test scores does not take into account where those students started the year. Value-added methods purport to isolate a teacher’s contribution to student learning in a year. However, the model specifications contain many important assumptions not well-understood by most policy makers, which can substantially affect estimates of teacher effectiveness (see McCaffrey et al. 2003). For example, the various models make different assumptions about the persistence of teacher effects in students’ future learning, the need to account for differences in teachers across schools, and the importance of accounting for the different backgrounds of students taught. Recent research has also called into question whether these models can sufficiently account for the sorting of students and teachers into classrooms to generate accurate causal estimates of teacher performance (Rothstein 2008).
Finally, value-added approaches are quite complicated and, as a result, teachers never learn why they receive the scores they do.

Beyond these methodological challenges, achievement tests also cover a limited array of content domains and are imprecise measures of student performance in those areas. Tests typically assess only a limited range of subjects and grades; thus, using statewide examinations in a pay-for-performance system may exclude early elementary teachers, high school teachers of upper grades, and teachers of specialty subjects. Tests are also quite noisy estimates of a student’s true performance. For example, a student’s test performance may reflect true learning or may simply be the idiosyncratic result of successful guessing (Koretz 2008). The effect of this substantial error should not be understated as it attenuates the relationship between teachers’ actual performance and estimates of their effectiveness, thereby limiting the incentive effects of any performance bonuses.

Parents and society expect schools to achieve a wide range of educational goals beyond successful test performance. They want students to grow up to be good citizens, creative thinkers, and productive members of society. They also want to see students graduate and, in many cases, go on to post-secondary education. Districts thus face a “multi-tasking” problem in designing performance incentives; they expect their employees to undertake a variety of tasks, but only have objective performance measures about some (Holmstrom and Milgrom 1991). In response, teachers are likely to focus additional attention on meeting the goals that are measured and reduce their attention on others. For example, they may teach to the test rather than pursuing a broader curriculum, encouraging creativity, or developing social responsibility (Koretz 2008). In the extreme, narrowly-tailored incentives can produce perverse responses, for example, when

---

6 As value-added models essentially examine how much student performance improves from one test to the next, the estimate of this growth can be particularly sensitive to the imprecision with which these tests measure proficiency.
schools or teachers encourage low-performing students to be absent or even cheat on the test (Jacob and Levitt 2003; Figlio and Getzler 2002). Thus, districts must provide incentives that directly reward the specific behaviors they want to encourage (Kerr 1975).

Given that complexity, some evidence suggests that subjective performance evaluation may prove a superior basis for making compensation decisions (Baker, Gibbons, and Murphy 1994). These professional evaluations, advocates contend, would combine many elements and enable a more holistic assessment. However, for such approaches to be successful, supervisors must have the capacity to make fair and nuanced decisions about teachers’ performance and teachers must view the evaluations as reflecting their true performance (MacLeod 2003). Furthermore, the threat of favoritism and personal bias entering into decisions about compensation remains a substantial challenge when using performance evaluations. Given that schools moved to the single salary scale to eliminate patronage and differential treatment, returning to a compensation system based on professional evaluation may simply lead to replaying the mistakes of the past.

Finally, districts can combine information from professional evaluations, students’ test performance, and other measures to make decisions about performance awards. This approach uses a wider range of information and provides incentives for teachers both to improve student test scores and to satisfy the requirements of a successful professional evaluation. However, using multiple measures also comes with challenges. Program designers must assign weights to different aspects of performance in making compensation decisions. Diluting a performance incentive by providing multiple avenues to success also reduces the motivational impact of any one element. Finally, combining multiple elements into a single program element may generate
resentment and resistance from teachers who produce high student achievement results but do not have sufficiently high professional evaluation marks to receive the bonus, or vice versa.

Currently, most pay-for-performance plans rely on standardized test scores rather than professional evaluation. Increasingly, new initiatives are using student growth or more complicated value-added models to identify top-performing teachers or schools. Several districts, such as Douglas County, Colorado use professional evaluations solely to award performance-based pay. Others use professional evaluation to move teachers along a career ladder, which results in higher compensation. Few districts incorporate multiple measures into a single program element, although as we discuss below, Hillsborough County, Florida, has implemented such a model.

(2) How to identify top performers: Relative rankings or standards-based?

Once a district has decided how to assess performance, it must determine which teachers have done well enough on this “performance” measure to deserve the reward. Here, two main approaches prevail: relative or standards-based determinations. Relative awards form the basis of most of the pay-for-performance literature in economics— the key concept is a rank-order tournament, which arrays employees according to their performance and then rewards individuals above a certain threshold. Alternately, districts can provide rewards to all teachers who reach a specified, external standard. Again, both choices have advantages and disadvantages.

Relative award plans provide several important advantages for schools and teachers. First, they insulate teachers from common “performance shocks,” policies or events that affect all teachers in the district. For example, if all students in a district struggle because the district introduces a new but ineffective curriculum, the best performing teachers would still receive a
bonus. Second, they serve as a commitment mechanism for the district: if the district agrees to reward 25% of teachers, it cannot set a standard so high that no teachers reach it. Third, these types of plans limit the district’s financial risk because it can allocate a set budget for the rewards.

On the other hand, relative plans may introduce unwanted competition among teachers, particularly if the relevant comparison group is narrowly defined. In other words, rewarding the top teachers in each school would produce more competition than rewarding the top teachers throughout the state. At the same time, paying a bonus only to the top teachers in the state might disproportionately reward teachers who work with already advantaged student populations. Thus, the comparison groups used to determine relative performance must be considered carefully—again, there are no easy answers.7

By rewarding all teachers whose performance meets an external standard, standards-based plans allow all teachers to earn bonuses simultaneously, thereby limiting the negative effects of competition. In fact, these plans may even encourage collaboration as all teachers try to boost their students’ performance. However, as discussed above, they leave teachers open to the risk of policies or events that may negatively affect everyone in the district, thereby preventing many teachers from getting the bonus. They also require careful attention to the standard-setting process because targets that are too rigorous will not seem attainable and therefore not encourage extra effort, while standards that are too easily reached will not lead to behavioral changes. Finally, standards-based approaches leave the district or state administering the program at substantial financial risk because all teachers can, in theory, meet the standard

---

7 Importantly, many value-added models implicitly compare certain types of teachers because of model specification decisions. For example, including school-fixed effects compares teachers within the same school, while including student-level covariates compares teachers with similar student populations. While these issues fall outside the scope of this study, they remain of real importance for teacher compensation reform based on such measures.
and earn the bonus. Thus, standards-based plans may distribute less total money to teachers when districts act to limit such financial risks.

Today, most pay-for-performance plans use some form of relative teacher rankings to allocate rewards. The difficulty of developing external standards appears to prevent many districts from implementing standards-based programs, but several districts have hybrid programs, using relative rankings for a broader area (e.g., the entire state) to identify top performers in the district. In such programs, all teachers in a district could, in theory, achieve the bonus if they all outperform similar classrooms in the state. As we discuss later, Charlotte-Mecklenburg employs such a model.

(3) At what level to provide rewards: Group or individual?

Districts must also decide whether to provide incentives and rewards for individual or group performance. Individual programs, which reward teachers for the performance of students in their classroom, seem to follow logically from the “egg crate” nature of schooling (Lortie, 1975). These plans allow more directed and powerful incentives for individual performance. Group rewards, which offer bonuses to entire instructional teams or schools for group-wide outcomes, acknowledge the technical difficulties in accurately identifying individual teacher’s contributions to student growth\(^8\) and can prevent some of the competition that may arise under individual plans.

Many actors contribute to student learning. Even in elementary schools, where the modal arrangement assigns one teacher to a single class of students for core subjects, students typically meet with different teachers for physical education, arts, and music. In middle and high schools, the challenges are even more substantial as social studies, science, and ELA teachers may all

---

\(^8\) Importantly, though, examining group-level or school-level value-added only resolves some technical challenges. Other sources of error can make estimates of group-level performance equally problematic.
contribute to a student’s reading performance. Providing group incentives eliminates the need to attribute student learning to a single teacher. It also allows teachers in non-tested subjects or grades to receive rewards and may prevent the competition inherent in individualized pay-for-performance. In fact, having all teachers work toward a common goal for all students may promote collaboration. Furthermore, Neal (2007) argues that group performance awards can create internal accountability mechanisms whereby teachers have incentives to identify and help struggling colleagues. This internal accountability may actually prove stronger in effecting positive changes in instructional practice than simple cash rewards.

However, group plans also produce several challenges. First, the incentives for teacher performance are much more diffuse than in individual plans; as the size of the group increases, the “free rider” problem becomes more severe. Because each individual teacher contributes only one small part to the overall group performance, the incentives for any additional effort are reduced. Thus, group awards may need to be larger than individual bonuses to effect the same change. Furthermore, group plans may introduce new organizational challenges. If schools do not have the appropriate structures and cultures to deal with internal accountability (Abelmann and Elmore 1999), holding teachers accountable for the performance of their colleagues may create resentment and counterproductive group dynamics. This issue may be particularly challenging if districts simultaneously seek to attract strong teachers to low-performing schools.

Both individual and group awards have received widespread attention nationally. Some districts implement both types of programs, often combining group and individual incentive components within a larger compensation strategy. In many cases, group rewards have proven more politically feasible than individual bonuses.
1.3 Pay-for-performance in practice

In the framework presented above, we provide a conceptual structure for the wide variety of pay-for-performance programs in use around the country. However, understanding how these programs actually function requires considering how they work in context. Here, we describe four prominent initiatives: Houston, Texas, Minneapolis, Minnesota, Charlotte-Mecklenburg, North Carolina, and Hillsborough County, Florida. We selected these large, urban districts because they have all been at the forefront of pay reform. They commit substantial amounts of money to performance-based bonuses and incorporate several elements into their compensation strategy. We have chosen to focus on district-level initiatives, rather than state programs, because even the most prescriptive state programs exhibit local variation. Each of these districts incorporates components from broader state programs that call for pay-for-performance, but they interpret and implement these elements in conjunction with their local needs. Furthermore, each district includes several different pay-for-performance programs in their compensation strategy, but they take quite different approaches to integrating these elements.

We gathered data on these programs from published reports, district Web sites, and phone interviews with key informants from the districts conducted between February and April 2008. In each example, we first explore the background and implementation of pay-for-performance. Then we address each program component in turn, explaining how the design used by the district reflects the decisions raised by our framework.

---

9 It is important to note that we have not focused on Denver, by many accounts the most dramatic and comprehensive approach to compensation reform for two reasons. First, unlike any other district initiative, ProComp essentially replaces the single salary scale. Thus, while its program elements fit neatly into our framework, the district’s overall strategy is so unique as to not be representative. Second, the Denver program has been widely studied and reported (see for example Gonring, Teske, and Jupp 2007).
Houston, Texas

The Houston Independent School District operates a host of pay-for-performance programs under its plan, ASPIRE—Accelerating Student Progress: Increasing Results and Expectations. All of these programs reward teachers for student achievement. One strand, ASPIRE III-B, offers standards-based, group awards based on state accountability ratings. The other three strands provide individual and group-based awards based on relative students performance (Figure B). The district also rewards teacher attendance, providing teachers an additional 10% of their ASPIRE bonus for perfect attendance and 5% if they miss fewer than two days.

- **District background**

Houston Independent School District, the seventh largest district in the country, serves over 200,000 students in 300 schools and employs more than 12,500 teachers. In 2005-06, Houston first implemented a wide-ranging pay-for-performance plan, now called ASPIRE, which incorporates several different elements. The rewards are substantial, with teachers being able to earn up to $8,000 through a combination of school-wide and individual-level incentives. In the program’s first year, approximately 60% of the teachers received some performance award, although many earned only small amounts. Houston spent approximately $23 million of local funds and $2.7 million from a federal Teacher Incentive Fund grant on the program in 2005-06. Teachers also qualify for incentive funds from the state of Texas through several award programs. The state spends approximately $200 million a year in incentives targeted to high-achieving schools with large percentages of low-income students. Beginning in 2008-09, the state plans to set aside an additional $150 million for the District Awards for Teacher Excellence,
which will provide funding to all local districts. Sixty percent of the funds must reward teachers for student academic growth or achievement.

When it was introduced, pay-for-performance in Houston encountered several serious problems. A glitch in the system overpaid bonuses to some part-time teachers and the district’s attempts to recover those funds led to negative publicity about the program. Following a request from the Houston Chronicle, the district also released the names and bonus amounts for all teachers in the district, resulting in substantial resentment. Finally, the initial value-added formula produced some inconsistent results, with highly respected schools not receiving bonuses. As a result, the district reconfigured the program, renaming it ASPIRE and contracting with William Sanders, an early developer of value-added methodology, to implement his EVAAS model in 2006-07.

While some teachers support the new plan, the implementation problems have turned many educators against it; some have even refused to accept their bonus money. Because teachers cannot understand or gain access to the proprietary EVAAS statistical model, some regard the program as a “lottery” because they do not know what they must do to achieve a bonus. For example, one teacher earned a $7,590 bonus one year but no award the previous year. According to the Houston Chronicle, “he can't explain why—he said he hasn't changed how he teaches—but he appreciates the extra cash” (Mellon 2008). Houston offers four primary pay-for-performance components.

- **ASPIRE Strand I: Value-Added Campus Improvement**

  This relative, school-level award uses value-added student achievement test scores to determine eligibility. Using Sanders’ EVAAS value-added model, the program determines a school-level composite score across all grades and subjects. This model measures student
improvement on the Texas Assessment of Knowledge and Skills (TAKS) and national norm-referenced achievement tests such as the Stanford-10. The district ranks schools at the elementary and secondary levels. At each level, the district provides all teachers in schools that rank in the top 25% a $1,000 bonus, while those in the next 25% earn $500. The maximum award for non-instructional staff is $500.

• **ASPIRE Strand II: Teacher Progress**

  The Teacher Progress strand offers individual, relative pay-for-performance for core academic subject teachers in grades three through eight, again using value-added student achievement test scores. Within each subject, the district ranks teachers; those in the top 25% earn a $5,000 bonus, while those in the next 25% earn $2,500. For teachers of multiple subjects, the district uses a sliding scale to determine the total payout such that a teacher in the top 25% on all subjects can earn the maximum $5,000 bonus. To avoid issues in non-tested grades, core teachers in early elementary or high schools can earn bonuses for school-wide value-added gains; early elementary teachers can earn up to $2,500, while high school teachers can earn the full $5,000 bonus. Non-core teachers are not eligible to receive bonuses under this component.

• **ASPIRE Strand III-A: Campus Improvement Awards**

  Schools that earn ratings of Academically Acceptable or higher under the Texas state accountability system are eligible for this relative, school-level award. Here, the district compares school-level student improvement in reading and math test scores to a group of 40 demographically similar schools across the state. Although this program does compare schools, in many ways it is not a pure relative system because much of the reference group comes from outside of the district. Teachers in schools that rank in the top 25% of these comparable schools earn a $500 bonus, while those in the next 25% earn $250.
• **ASPIRE Strand III-B: Campus Achievement Awards**

This school-level, standards-based award pays all teachers in schools rated as Exemplary or Recognized under the Texas state accountability system a bonus of $300 or $150, respectively.

• **Summary**

Houston’s ASPIRE program represents an extensive pay-for-performance effort, offering substantial relative and standards-based awards at both the individual and school level. All of the district’s programs focus on student achievement demonstrated on standardized tests. Houston has invested significant amounts of local funding in the effort, supplemented with state and federal grants.

By uniting these programs under the ASPIRE umbrella, the district has created what it calls a “cohesive educational-improvement model.” The effects of this integration, though, remain unclear. By combining these types of programs, the district may limit some of the competition produced by a purely relative, individual approach. However, a large portion of the award money available through ASPIRE does come in these individual bonuses. Furthermore, in many ways, the program operates as a set of distinct options that offer teachers additional pay. The individual rewards under Strand II appear to be the most controversial because of substantial questions surrounding the methodology. If teachers continue to perceive that the awards do not relate to instructional quality or effort, then the program will likely not provide sufficient incentives to alter instructional practice. Hillsborough County, Florida faces similar challenges in its pay-for-performance programs.
Hillsborough County, Florida

Hillsborough County School District has created a local pay-for-performance plan and participates in the state’s Merit Award Program (MAP). The local plan, which preceded the state program, provides an individual salary supplement based on a standards-based professional evaluation. Under the more recent state-sponsored Merit Award Program, the district uses mixed measures, incorporating both professional evaluation and student achievement test growth to pay individual teachers for relative performance. Finally, the district rewards individual Advanced Placement and International Baccalaureate educators for student performance on the standards-based course tests (Figure C).

• District background

Hillsborough County School District serves nearly 200,000 students in West-central Florida, including the city of Tampa. The district began working towards a pay-for-performance plan around 2000 with a joint effort by the local teachers union and district administration. In 2001, the Florida legislature passed E-Comp, which mandated that districts implement performance-based pay. E-Comp proved to be very unpopular with Florida teachers and school districts. In fact, many districts deliberately created plans so rigorous that no teachers could achieve the performance bonus, thereby preserving the status quo.

Public outcries and lack of buy-in from local districts led the state to abandon the program and develop Special Teachers Are Rewarded (STAR) in 2006; it funded STAR with nearly $150 million for 2006-07. This program again called for districts to implement performance-based pay, but it accompanied this mandate with substantial funding if they completed a state-approved plan. The STAR requirements mandated awards of 5% to the top 25% of teachers. Because Hillsborough County already had pay-for-performance, the district
quickly wrote a new proposal, based on their experience with their current program, and it was approved. In March 2007, however, the state abandoned STAR in favor of a more flexible Merit Award Program. Hillsborough County then revised its STAR plan and became the first district in the state to achieve approval under MAP.

Despite the labor-management collaboration throughout the process in Hillsborough County, many teachers reported dissatisfaction during the program’s first year, suggesting that the plan did not reward the right teachers. For example, half of the district’s Teacher of the Year finalists failed to earn the bonus (Stein 2008). Furthermore, according to district officials, teachers perceived that those who worked in wealthier schools were more likely to earn bonuses. Independent analyses revealed some evidence of this problem: the St. Petersburg Times reported that “almost three-fourths of the nearly 5,000 teachers who received merit pay worked at the county's more affluent campuses” (Stein 2008). To respond to these issues, the administration and teachers union held meetings with teachers and began to retool the plan. Still, though, the teachers union president sees the bonuses as a “lottery,” with rewards not clearly dependent upon instructional performance (Stein 2008).

- **Merit Award Program**

  This individual, relative plan rewards high-performing teachers on the basis of professional evaluation and student test scores. Principals provide teachers with a numerical score from classroom evaluations using a detailed rubric that allocates points based on ratings in specific performance categories. Teachers need to earn at least 123 of 144 possible points on the performance appraisal and have no “unsatisfactory” ratings to be eligible for a merit award. Every teacher also gets an effectiveness score based on student performance growth. The district uses the Florida Comprehensive Assessment Test (FCAT) when possible, but allows national
achievement tests and district assessments to be used in other grades and subjects. The district has developed detailed “value tables” that compare a student’s performance “level” on the pre-test and the post-test and allocate points accordingly. Teachers get some points for students who stay at the same level (above the lowest category) and more points for moving students across levels. Teachers lose points for any students who drop levels.

The district combines the performance appraisal (40%) and effectiveness score (60%) to create a total teacher score; this weighting structure derives from the state’s MAP guidelines. It then ranks all teachers in separate groups defined by their specific subject-area (e.g., reading, math, science, etc.) and grade level (i.e., elementary, middle, and high school). In all cases, teachers in Title I and non-Title I schools are grouped separately. The district ranks teachers of multiple subjects in all possible groups and uses their highest ranking to determine award payments. For personnel who serve school-wide roles, the district uses school-wide reading and math test averages. In each group, the district provides an award of 5% of the average teacher salary to the top performing teachers until the MAP funding runs out. In 2006-07, approximately 37% of teachers in each group earned a MAP award of $2,100.

- **Local pay-for-performance plan**

Hillsborough County also operates a standards-based, individual plan that uses professional evaluations to reward high-performing teachers. Teachers volunteer to enter the process. To receive the 5% salary supplement, teachers must demonstrate through a portfolio and a conference with their building principal that they have performed at an exceptional level. Because the portfolio process is time-consuming, many teachers do not apply for this award. In 2006-07, approximately 1,500 of the district’s 15,000 teachers received this bonus. The district
finances these awards through a pay-for-performance trust fund that it developed to insulate the
program from local financial pressures.

- **Awards for Advancement Placement/International Baccalaureate Performance**

  The district provides additional standards-based, individual performance pay for teachers
of Advanced Placement (AP) or International Baccalaureate (IB) courses whose students
perform well on the final exams. For each student who earns above three on the AP examination
or above four on the IB test, the teacher gets $50 up to a maximum of $2,000—or $4,000 for
teachers of both AP and IB courses. All AP and IB teachers in the district can earn this award if
their students meet the pre-set standard.

- **Summary**

  Hillsborough County’s two main pay-for-performance plans take somewhat unusual
approaches to measuring teacher performance and providing individual awards. The local,
standards-based plan uses a rigorous evaluation and portfolio process, which limits participation
to teachers who choose to invest the additional effort required. The MAP program uses multiple
performance measures, weighting student test score gains and professional evaluations to
determine relative awards. By using mixed measures, the MAP program mitigates some of the
challenges faced by pure achievement-based efforts. It encourages teachers not only to raise their
students’ test scores, but also to improve their instructional practices. In theory, combining
relative and standards-based components can moderate some of the competition that arises.
However, the MAP program has received substantial publicity and has become seen as the
primary program. Furthermore, methodological issues with the program have produced real
incentive problems because many teachers see their bonuses as decoupled from actual classroom
performance.
The district very consciously operates its different pay-for-performance programs separately, even keeping the funding streams entirely distinct. The district adopted a MAP plan to take advantage of the generous state financing; while the district and teachers both seem to have real concerns with the program’s methodology, they see the program as the only way currently available to introduce additional monies into teacher compensation in the state of Florida. This approach of adding on disparate plans as new opportunities present themselves mirrors the experiences in Charlotte-Mecklenburg.

**Charlotte-Mecklenburg, North Carolina**

Charlotte-Mecklenburg operates several local and state-run *standards-based* award programs. Three of the programs focus primarily on *student achievement*, while the district’s Retention Initiative incorporates other measures. Both the ABC and local accountability bonus offer *group-based* awards for school performance, while the other two programs provide bonuses for *individual* teachers (Figure D).

- **District background**

  Charlotte-Mecklenburg Schools (CMS) is the nation’s 20th largest district, serving 132,000 students in the Charlotte area. CMS was an early adopter in the most recent pay-for-performance wave, implementing rewards for principals in the mid-1990s based on student test scores. Around the same time, the state began its ABCs of Public Education accountability program, which included school-wide performance bonuses. North Carolina first piloted the program in 10 districts and expanded it to the entire state in 1996-97.
Recently, Charlotte-Mecklenburg began providing rewards for individual teachers. Like the state program, CMS began with a pilot program in 11 schools. Having concluded that the pilot was succeeding, the district extended the program to include all 65 FOCUS schools, low-performing schools that receive additional resources. The district has recently undergone some transition in its teacher compensation programs. It received a $1.9 million Teacher Incentive Fund grant in 2007 and plans to use this federal money to help implement a broader pay-for-performance program. As a result, the district ended the pay-for-performance pilot in 2006-07 and plans to stop the STAR program in the FOCUS schools after 2007-08. Informed by the pilot program, the district plans to unveil a new pay-for-performance program in 2008-09 that uses student achievement and classroom evaluations to reward teachers. It has also contracted with the Community Training and Assistance Center (CTAC), the same organization involved in Denver’s ProComp program, for advice.

- **North Carolina state ABC’s of public education**

This statewide accountability plan offers school-wide standards-based\(^{10}\) awards for student test score growth. The state predicts how much growth a student should make each year based on the average growth shown in the previous two years; the difference in actual and expected performance is the student’s “academic change.” For elementary schools, the state averages student-level academic change scores and, if the scores are greater than zero, determines that the school met “Expected Growth.” In high schools, other indicators such as dropout rate and changes in the percentage of students taking college preparatory courses factor into the rating. Schools in which the ratio of students who meet the academic change standard to those who fail it exceeds 1.5 are labeled “High Academic Change” school. Teachers in these

\(^{10}\) While the state formula essentially compares schools to the average performance in the state, a relative measure, from the perspective of the district, this award operates as a standards-based program because all schools can earn the bonus if they exceed the external, state-defined standard.
schools earn a $1,500 bonus, while teachers in “Expected Growth Schools” earn $750. All certified teachers in the school receive the bonus, regardless of subject area.

• **Local accountability bonus**

  CMS offers an additional school-wide, standards-based award for student test scores to provide teachers incentives to reach state and federal accountability goals. The district divides award recipients into Tier 1 and Tier 2. Tier 1 schools meet Adequate Yearly Progress in all subgroups and earn either “High Academic Change” or “Expected Growth” status in all subgroups under the North Carolina accountability system. Tier 2 schools reach one of these goals but not the other.

  This program is standards-based and, in theory, all schools can meet the standard. To minimize the financial risk, the district sets aside a fixed budget for the program, which it divides among all schools that meet the standard. Thus, the award amount varies from year to year and is relative—the more schools that win the award, the smaller the bonuses. Teachers in Tier 1 schools earn a larger bonus. All employees assigned to a single school site are eligible, with the exception of teachers who have “documented unsatisfactory performance.”

• **Successful Teacher Administrator Reward (STAR)**

  This standards-based individual bonus rewards teachers (and building administrators) for high academic growth or achievement levels on statewide end-of-grade or end-of-course assessments. The teacher-level award is only available for teachers of core, tested subjects in the district’s 65 FOCUS schools. Here, the district calculates student performance using a formula similar to the state’s, predicting students’ scores based on their growth over the previous two years. To receive the bonus, a teacher must have 60% of students making “high growth.” For teachers whose students do not have end-of-grade or end-of-course growth scores, the district
uses performance levels, counting the percentage of students who score at levels three or four on the accountability system. Eligible teachers earned a $1,400 bonus in 2006-07.

- **Annual Retention Incentive for Master Teachers**

Finally, the district offers a retention incentive for master teachers to remain in one of the district’s 65 FOCUS schools. To earn this standards-based individual award, teachers must have consistent “at standard” evaluations and principals’ recommendations must report that they practice “cultural sensitivity.” Eligible teachers can earn the reward for meeting a variety of criteria, one of which is standards-based student performance on the state standardized tests. Other possible criteria include NBPTS certification, licensure in multiple subject areas, or Advanced Placement, Academically Gifted, or International Baccalaureate credentials. Teachers with at least four years of service and an advanced degree earn a Level I award, worth $2,500. Teachers with eight or more years of experience who are enrolled in a graduate program but have not yet completed their degrees earn $1,500 for a Level II award.

- **Summary**

Charlotte-Mecklenburg operates several distinct programs offering teachers standards-based rewards primarily for student achievement growth at both the individual and school-wide levels. The district has capitalized on a state context that provides sufficient data systems and analytic capacity to evaluate student learning gains for use in the pay system. It also participates in a statewide group award program that operates entirely outside of district control, unlike the MAP program in Florida, which requires individual districts to develop plans. These programs provide incentives for teachers to improve their students’ test performance and to work together with other teachers in the school to raise overall student scores.
The district’s local programs arose from the general goal of developing broader compensation reforms. CMS has added separate program components in attempts to remedy specific problems or as additional funding has become available. For example, recognizing that it was losing high-performing teachers from its lowest-performing schools, the district developed a series of rewards to encourage teachers to remain in those schools. The piecemeal planning of these programs has left the district with a set of largely independent programs. However, the district has begun efforts to develop a more integrated compensation strategy using its new Teacher Incentive Fund grant; in this regard, the district may be moving more toward a model like that in Minneapolis, which attempts to unite all pay-for-performance elements under a single program structure.

**Minneapolis, Minnesota**

Minneapolis has incorporated its pay-for-performance awards into the Alternative Teacher Professional Pay System (ATPPS). Here, the district offers two *relative, group-based* awards: the MnTAP School Bonus for *student achievement* and the Quality Performance Award for *mixed measures*, including student performance and a variety of other indicators. The district also provides an *individual, standards-based* bonus for *professional evaluation* in its Teacher Advancement Program schools (Figure E).

- **District background**

  The Minneapolis Public School District is the largest in the state of Minnesota, serving 35,000 students. Minneapolis began pursuing pay-for-performance in 2002, when the district and the Minneapolis Federation of Teachers received a grant from the state to develop an alternative
compensation system (ProPay). In 2004, the district also implemented the Teacher Advancement Program (TAP), a comprehensive approach to school staffing and professional development that includes performance-based pay, in several schools. TAP is a national effort, developed by the Milken Family Foundation and now operated by the National Institute for Excellence in Teaching. These actions helped spark the state, in 2005, to pass the ATPPS legislation, which provided state funding for alternative teacher compensation. In Minneapolis, the union and management collaborated to develop a plan that incorporated both the ProPay and TAP systems under ATPPS, which the district first used in 2006-07.

With ATPPS, the district replaced the traditional salary scale with a similar structure that moves teachers along lanes for accruing Professional Growth Credits (PGS) rather than simply post-secondary education credits. Teachers move across one lane for each 12 PGCs earned, with a permanent salary increase of $1,000. In any year, a teacher can move no more than two lanes. Teachers still move up salary steps with each year of experience. Teachers can earn PGCs for a wide range of activities, such as earning course credits, undertaking professional development, adopting special responsibilities, teaching in hard-to-staff schools, obtaining a license in a high-need subject area, or improving student achievement. Thus, the district incorporates many alternative compensation components, including pay-for-performance, into the ATPPS program. Here, we address several elements that offer teachers bonuses for their classroom performance.

- **Quality Performance Awards**

  Teachers can earn PGCs for school-level performance on multiple measures. The district scores schools in 33 Quality Performance indicators. Many of the indicators address student test score growth, on both state and nationally normed tests. Other indicators include student attendance, measures of instructional equity, student discipline, and results from student and staff
satisfaction surveys. Most of these indicators are standards-based, although the district does include several relative value-added measures that compare the schools to their “expected” growth.

Schools earn a score from one to five on each of the 33 indicators, and the district then averages these scores to create a final composite Quality Performance Award score. Teachers in the school earn a PGC for each point. Thus, teachers in top performing schools earn 5 PGCs, while those in the lowest performing schools earn 1 PGC.

- *MnTAP Teacher Advancement Program*

  Minneapolis also participates in the Teacher Advancement Program (MnTAP), which provides individual incentives for standards-based professional evaluation and school-wide relative bonuses for student test score growth. Currently, 15 schools, representing nearly 25% of Minneapolis teachers, use the program.

  All MnTAP teachers receive a Skills, Knowledge, and Responsibility (SKR) score—from one to five—based on professional evaluations done by the mentor, instructional coach, and building administrator. According to the MnTAP documentation, “a score of three is considered solid-proficient teaching where learners will make good progress and expected growth” (MnTAP Implementation 2008, 4). Each of these three evaluations counts for 30% of the total score; a self-assessment accounts for the remaining 10%. Teachers receive three PGCs for every final score point. Thus, a teacher with an evaluation score of four would earn 12 PGCs, worth a permanent salary increase of $1,000 per year.

  Teachers in MnTAP schools are also eligible for the Quality Performance Award bonuses (up to five PGCs). In addition, they can earn up to three PGCs for relative school-level value-
added measures in both reading and math (up to six PGCs total). Here, teachers in schools that make greater than expected growth in mathematics and/or reading earn the bonus.

- **Summary**

  Minneapolis rewards teachers for performance in a variety of dimensions. The district offers relative group awards for rankings in student achievement growth and multiple measures, while its individual plan, available only in TAP schools, provides bonuses based on professional evaluations. By using student test scores, professional evaluations, and other measures, Minneapolis provides incentives for teachers to improve student outcomes in a variety of ways. However, this diversity makes the incentives to increase student achievement less powerful.

  Unlike many other states, Minnesota developed its pay-for-performance program largely in response to activities already underway in Minneapolis. Minneapolis is also the only district we examined that provides permanent salary increases for its performance awards, rather than one-time bonuses; this practice remains relatively rare across the country. While providing a permanent increase puts substantially more money behind the reward and eliminates the penalty of having teachers’ salaries decrease from one year to the next, its incentive effects depend in large part on teachers’ career expectations.

  Furthermore, Minneapolis’s ATPPS program represents the most integrated approach to teacher compensation among the districts that we consider here. Like Houston’s ASPIRE program, ATPPS combines several elements into a single pay-for-performance plan. Minneapolis united several disparate elements, including the pre-existing ProPay 1-2-3 and MnTAP as well as bonuses for differentiated roles and for teaching in hard-to-staff schools, into a single ATPPS framework. In fact, Minneapolis has integrated these pay-for-performance
elements within its entire teacher salary scale. It remains to be seen if such changes represent a symbolic reorganization or a systematic alignment of pay components with district strategies.

**Conclusion**

This study has outlined three key areas in which pay-for-performance programs differ: how districts define performance, how they identify top performers, and at what level they provide awards. District can choose to use student achievement (as in Houston), professional evaluations (as in Minneapolis’s SKR bonus), or multiple measures (as in Hillsborough County’s MAP plan) to assess teacher performance. They can identify which teachers deserve the awards through relative rankings (as in most of Houston’s programs) or comparisons to an external standard (as in Charlotte). Finally, districts can choose to provide awards to individuals (as in Hillsborough County) or entire groups of teachers (as in Minneapolis’s QPA program). Districts also make different choices in combining these dimensions; for example, Houston uses relative rankings to determine their primary school-based award under ASPIRE, while Charlotte-Mecklenburg uses standards-based measures for their school-wide bonuses. **Figure F** summarizes all of the different elements described above.

In addition to the primary design decisions outlined above, policy makers also face a wide range of other questions surrounding program design and implementation. For example, should the program operate at the state or local level? How much should it provide in performance awards? Should it pay all rewarded teachers a constant amount or use a sliding scale based on performance? Should it reward teacher performance directly through salary supplements or more indirectly through promotions along a career ladder? Should it allow veteran members to opt in or require participation from all teachers? Should it provide awards as
one-time salary bonuses or as permanent salary increases? All of these additional choices introduce important trade-offs and have limitations; they are not trivial. For example, providing awards as permanent salary increases may be more effective at encouraging teachers to undertake the capacity-building professional development necessary to improve their practice. But, the long-term commitment will likely be more costly for the district and may dilute the incentives in any given year. The incentive effects will also vary according to each teacher’s career expectations: teachers who plan to stay in teaching for a long time will see a permanent increase as relatively more valuable as teachers who plan to leave the district.

Furthermore, many of these decisions also have important implications for the political challenges of implementing and sustaining pay-for-performance. This study has focused largely on the technical issues surrounding design choices. However, the political barriers in implementing any compensation reform remain large. Simply developing a well-designed plan is not sufficient. Instead, districts looking to revise their teacher pay systems must think carefully about the process. Many of the most comprehensive—and arguably most successful—pay-for-performance initiatives have been developed jointly by representatives of the teachers union and district administration. Program designers have also learned from the pitfalls of the past. For example, they have worked to develop strong and stable funding sources that build confidence in the program’s ability to maintain its commitments. Although these political issues fall outside the scope of this study, policy makers and districts must acknowledge their importance.

The wide variation in pay-for-performance programs across the country suggests that the correct design for any district depends in large part on the local context, including the organizational and political culture. In many districts, deeply rooted school culture influences the reform agenda. For example, a district with a strong culture of internal accountability may prefer
a group-based plan, while a district without such a culture may choose to provide individual rewards. Furthermore, many other aspects of the local context, including the relationship between labor and management, the history of collaboration or competition in the district, the characteristics of the teaching force, the nature of the local labor market, the technical capacities of both teacher evaluation and student testing systems, and the personalities of local political actors can all play an important role in determining the success of any program. The same program that operates effectively and provides clear incentives that align with district goals in one setting may backfire and draw intense opposition in another. Gonring, Teske, and Jupp (2007), who worked on compensation reform in Denver, argue that districts should not import ProComp wholesale. In fact, they “cannot imagine a more dangerous undertaking, as [they] believe it is necessary for teachers to help build and sustain a system that works for them” (Gonring, Teske, and Jupp 2007, 136). Districts must design program elements that promote the appropriate incentives and that are practically viable in the local context.

In addition to designing individual programs, districts—including all four that we examined—often combine multiple elements into a more comprehensive pay-for-performance plan. This diversification arises for several reasons. First, districts may decide to develop multiple opportunities for teachers to earn awards in order to limit the negative reactions to these plans and moderate the unintended consequences of making choices about these dimensions. They may also pursue multiple pay-for-performance programs opportunistically to secure available funds from various sources.

Diversification can help make a program politically feasible by satisfying multiple stakeholders. Pay-for-performance programs are controversial, often coming under fire because they use student achievement data to measure teachers’ effectiveness. In these settings, including
multiple opportunities for teachers to earn performance bonuses may be a realistic strategy given
the serious limitations with value-added methodologies. Although value-added approaches have
improved substantially over the past decade, serious questions remain about how well they
account for uneven student learning trajectories and the nonrandom assignment of students into
classrooms. Problems with the methodology can produce violations of face validity. In some
cases, tests reveal that teachers widely recognized as being exceptional actually produce little
student learning and are not effective in the classroom. However, when large numbers of
teachers whose students are seen to do well do not earn rewards, the pay-for-performance
program can have trouble achieving legitimacy. Several districts have struggled with their award
formulas to such an extent that teachers of the year and other recognized educators have failed to
qualify. Such misalignment can raise serious doubts about a plan’s success and lead teachers to
see it as a lottery, which they might win but over which they have little control. If teachers see
the award as unrelated to any effort or ability, even substantial bonuses will not promote
productive behaviors.

Including multiple opportunities for teachers to earn rewards can also eliminate some of
the perverse responses to the incentives, such as teaching to the test, inherent in narrowly-
tailored compensation plans. Here, diversification can serve as a possible remedy to the multi-
tasking problem in education by providing rewards for teachers to undertake a wide range of
important behaviors. In this regard, some districts attempt to integrate multiple elements into a
more comprehensive compensation strategy. In an effort to create incentives that focus teachers
on the behaviors most valued by the district, a comprehensive and thoughtfully designed strategy
that combines multiple elements may have the greatest chance for success.
Finally, many districts add on separate components as political and financial opportunities arise. For example, Hillsborough County had an existing pay-for-performance plan in place when the state of Florida offered additional funds for any district through the Merit Award Program; to take advantage of the money on the table, the district developed a plan that met the state’s requirements. Thus, the plan did not reflect the district’s preferred choices about program design because it needed to operate within state-defined constraints. Not surprisingly, while MAP complements the district’s local pay-for-performance program, the two initiatives continue to operate separately.

As a result, programs may not follow a coherent approach to teacher compensation that aligns with the district’s instructional goals. While each individual component may contain incentives that drive teacher behavior, these incentives may not produce instructional changes that work in unison. Because pay-for-performance programs typically entail substantial investments, districts should consider deliberately integrating their approaches to pay reform so that they provide clear incentives, which are aligned with their priorities for improvement.
Figure A. A framework of pay-for-performance programs.

<table>
<thead>
<tr>
<th>How to identify top performers?</th>
<th>How to measure performance</th>
<th>At what level to provide awards?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Rankings</td>
<td></td>
<td>Individual</td>
</tr>
<tr>
<td>Standards-based</td>
<td></td>
<td>Group</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How to identify top performers?</th>
<th>How to measure performance</th>
<th>At what level to provide awards?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Rankings</td>
<td>ASPIRE II</td>
<td>Individual</td>
</tr>
<tr>
<td></td>
<td>ASPIRE I</td>
<td>Group</td>
</tr>
<tr>
<td></td>
<td>ASPIRE III-A</td>
<td></td>
</tr>
<tr>
<td>Standards-based</td>
<td>ASPIRE III-B</td>
<td>Individual</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Group</td>
</tr>
</tbody>
</table>

Figure B. Houston pay-for-performance programs.
Figure C. Hillsborough County pay-for-performance programs.

<table>
<thead>
<tr>
<th>How to identify top performers?</th>
<th>Student achievement</th>
<th>Professional evaluation</th>
<th>Mixed measures</th>
<th>At what level to provide awards?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Rankings</td>
<td></td>
<td></td>
<td>MAP</td>
<td>Individual</td>
</tr>
<tr>
<td>Standards-based</td>
<td>AP/IB Award</td>
<td>Local Plan</td>
<td>Individual</td>
<td>Group</td>
</tr>
</tbody>
</table>

Figure D. Charlotte-Mecklenburg pay-for-performance programs.

<table>
<thead>
<tr>
<th>How to identify top performers?</th>
<th>Student achievement</th>
<th>Professional evaluation</th>
<th>Mixed measures</th>
<th>At what level to provide awards?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Rankings</td>
<td></td>
<td></td>
<td>Individual</td>
<td>Group</td>
</tr>
<tr>
<td>Standards-based</td>
<td>STAR</td>
<td>Retention Incentive</td>
<td>Individual</td>
<td>Group</td>
</tr>
<tr>
<td></td>
<td>ABC</td>
<td>Local bonus</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure E. Minneapolis pay-for-performance programs.

<table>
<thead>
<tr>
<th>How to measure performance</th>
<th>Student achievement</th>
<th>Professional evaluation</th>
<th>Mixed measures</th>
<th>At what level to provide awards?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Rankings</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MnTAP – School bonus</td>
<td></td>
<td></td>
<td>Individual</td>
<td></td>
</tr>
<tr>
<td>Standards-based</td>
<td></td>
<td>MnTAP – SKR Bonus</td>
<td>Individual</td>
<td>Group</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QPA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure F. Overview of teacher pay-for-performance plan elements.

<table>
<thead>
<tr>
<th>Type of Program Component</th>
<th>Houston</th>
<th>Hillsborough County</th>
<th>Minneapolis</th>
<th>Charlotte-Mecklenburg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual, Relative Awards</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student Achievement</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed Measures</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Individual, Standards-Based Awards</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student Achievement</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Professional Evaluation</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Mixed Measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group, Relative Awards</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Student Achievement</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Professional Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed Measures</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Group, Standards-Based Awards</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student Achievement</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Professional Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed Measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
REFERENCES


