# CECR

# Emerging Issues

## Performance-Based Compensation Structures:

Considerations for Individual, Group, and Hybrid Programs

# Performance-Based Compensation Structures:
## Considerations for Individual, Group, and Hybrid Programs

By Amy Potemski and Cortney Rowland, American Institutes for Research;
Peter Witham, University of Wisconsin–Madison, Wisconsin Center for Education Research

## *Introduction*

A significant number of educator compensation reform efforts are under way throughout the country. These school-, district-, and state-level programs come in all shapes and sizes—some are small and focus only on a cohort of teachers or schools, whereas others are large and target entire districts or groups of districts. The structure of these compensation incentives varies substantially across these reform programs, ranging from individual educators who receive awards for their performance to rewards based solely on group performance. This variation in the structure of performance-based compensation programs is one of the most nuanced, least understood, and rarely discussed aspects of compensation reform efforts. The purpose of this paper, therefore, is to assist the stakeholders in alternative compensation initiatives to better understand the ways in which they can structure performance incentives in the context of their intended outcomes, local culture, and available options for measuring educator and student performance. With this information, stakeholders can select the reward structure most suitable for the needs of their specific education system.

This Center for Educator Compensation Reform (CECR) Emerging Issues paper focuses on the *structure* of performance-based compensation reform programs—who receive awards and how they receive them. Decisions about program structure generally address the following three considerations:

- The **unit of accountability** focuses on whose performance the school or district measures to determine the award (Heyburn, Lewis, & Ritter, 2010; Organisation for Economic Cooperation and Development, 2009). The unit of accountability can be an individual, a group, or an entire school. This report focuses on three units of accountability: (1) group, (2) individual, and (3) hybrid. Hybrid programs are those that reward both individuals and groups through the same program. Decisions about the unit of accountability a program uses often depend on local culture and program goals.

- **Measures of performance** are input and output measures states use to evaluate and reward the unit of accountability's performance. Measures of educator performance can include student performance on tests or other aspects

of student work (output measure), classroom instruction (output measure), or teacher characteristics (input measure). Deciding which measures of performance to use can be particularly challenging for teachers of grades and subjects for which most state and district standardized examinations have no tests.

- **Incentive eligibility** refers to two distinct characteristics of a performance-based reward program. The first characteristic is the scope of employees covered by the compensation program. For example, will teachers, principals, and/or other educational support staff at the school level be eligible to receive rewards? The second characteristic is whether programs limit performance awards to a fixed number of educators or offer all educators eligibility to receive an award so long as they meet a certain performance threshold. As with determining the unit of accountability, deciding which educators will be eligible for program incentives often depends on local culture and overall program goals.

Structuring and implementing a performance-based compensation program to achieve the maximum desired impact on instructional quality, teacher recruitment and retention, and student achievement is a complex process. Structuring the program in a way that fits the local context and preferred goals is crucial to the effective design, implementation, and sustainability of the effort. This Emerging Issues paper describes the research-based pros and cons of different reward structures and the considerations described above. The paper also provides several policy considerations for schools, districts, and states that are making decisions about the structure of their performance-based compensation programs. This paper serves as background information and a conversation starter for district- and state-level compensation reform committees that are making decisions about their own compensation programs.

Specifically, the paper organizes the discussion into three main sections:

- What do different performance-based compensation program structures look like? This section provides examples of several current programs that exemplify a variety of program structures.

- What does the research say about various types of program structures? This portion includes a discussion of perceived strengths and weaknesses of group versus individual awards, and the research on stakeholder attitudes and opinions about each structure.

- What are some of the key elements of deciding which structure works best given a particular context? For example, do certain data systems or school/district cultures lend themselves well to an individual-based program versus a group-based program?



*Deciding which educators will be eligible for program incentives often depends on local culture and overall program goals.*

## Types of Performance-Based Compensation Program Structures

As the field of alternative compensation programs has evolved, schools, districts, and states have experimented with a variety of program types or structures. The following examples of current performance-based compensation programs represent different units of accountability, including those that reward groups, individuals, and/or a mixture of the two, for their performance:

- **Group awards** provide incentives to all teachers based on student achievement gains at the grade or subject level, to teams of teachers, or to designated school-based personnel (e.g., library staff) at the schoolwide level.

- **Individual awards** recognize teachers based solely on what happens in their own classrooms. Generally, this reward structure takes the form of teacher pay for student achievement gains at the classroom level; however, incentives for teacher performance in the classroom or for activities that increase teachers' knowledge and skills could also be included under individual awards. For example, professional development activities, mentoring, and observation-based evaluations of teacher performance are all potential measures for individual awards.

- **Hybrid programs** provide awards based on multiple units of accountability. This reward structure usually provides incentives to all teachers based on schoolwide student achievement supplemented by individual awards based on classroom student achievement gains. Many compensation reforms take on a hybrid reward structure to balance out the weaknesses of each design with the strengths of the other. There is no one model that serves as an ideal hybrid program; instead, hybrid programs can take on a variety of formats.

## Examples of Programs That Provide Awards Based on Group Performance

The New York City Department of Education, with support from The Broad Foundation, the Fund for Public Schools, and the Partnership for New York City, funds a schoolwide incentive program for high-need schools.[1,2] The NYC Department of Education invited 201 high-need schools to participate; 199 accepted. The schools range from elementary schools to high schools. At least 55 percent of the union-represented school staff and the principal must vote in favor of the program for participation. The program provides bonuses of up to $3,000 per full-time United Federation of Teachers (UFT) member to schools that meet all of their annual performance requirements. If a school meets at least 75 percent of its performance targets, it is eligible to receive a bonus of $1,500 per full-time UFT member.

Similarly, the Texas Governor's Educator Excellence Grant program (GEEG),[3] allowed schools to tailor their reward structures to their particular contexts and needs. In a three-year pilot program, the GEEG targeted schools in the state's highest need areas and required schools to design their compensation models to meet two specific criteria: the programs must award bonuses based on (1) measures of student outcomes and (2) teacher collaboration. The state stipulated minimum and maximum amounts for the individual teacher awards, but otherwise, it left program details to the schools' discretion.

---

1   A link to CECR's map profile of this program can be found here: http://www.cecr.ed.gov/map/newYork.cfm
2   A link to CECR's case summary of this program can be found here: http://www.cecr.ed.gov/TIFgrantees/resources/caseSummaries.cfm
3   A link to CECR's map profile of this program can be found here: http://www.cecr.ed.gov/map/texas.cfm and a link to CECR's case summary of this program can be found here: http://www.cecr.ed.gov/TIFgrantees/resources/caseSummaries.cfm

## Examples of Programs That Provide Awards Based on Individual Performance

There are few examples of programs that base compensation decisions solely on individual teacher performance and even fewer that focus only on individual performance as measured by student achievement. In fact, programs that focus on individual performance primarily measured by student outcomes frequently face pushback from key stakeholders. One example of this is the history of performance-based compensation in Florida. As CECR details in a previous report (Max, 2007), Florida attempted to implement an initial performance-based pay program mandated by the state. Each district program relied solely on student learning gains at the classroom level. This approach faced strong opposition from districts and teachers because of its reliance on student test scores to measure individual teacher performance. To address the issue, Florida revised program requirements through two program iterations, STAR and finally the MAP program, that defined performance through student learning gains and principal evaluations of teacher performance. The MAP program also added teacher teams as an eligible award level. The use of individual and team award levels as well as multiple measures led to increased support by stakeholders in the state (Max, 2007).

Another example of a compensation program based specifically on individual performance is the Portland Professional Learning Based Salary Schedule (PLBSS) in Maine. Portland Public Schools and the Portland Educators' Association developed PLBSS based on the belief that the greatest predictor of student achievement is teacher learning. The program allows teachers to earn increases by accumulating Salary Contact Hours; 225 approved hours allows a Portland teacher to advance one lane on the PLBSS salary schedule. The implementation of this program resulted in much higher starting salaries for teachers. A new teacher with a bachelor's degree and no experience can now reach the top of the salary scale in 22 years. Using the former scale, it would have taken that same teacher 30 years. With teachers in charge of their own promotion and raises, they are more likely to accept the program.

## Examples of Hybrid Programs

Most hybrid programs include individual- and school-level data. Programs use team-level awards less frequently (Heyburn et al., 2010). A majority of Cohort I and II Teacher Incentive Fund (TIF) grantees (26 out of 33) implement a program in which all eligible educators receive awards based on school-level data or a hybrid model that included one or more of the following components:

- Teachers in tested subjects and grades receive bonuses based on classroom-level student growth, whereas teachers in nontested grades and subjects receive bonuses based on schoolwide student growth.

- Teachers in tested subjects and grades receive bonuses based on both classroom and schoolwide student growth, whereas teachers in nontested grades and subjects receive bonuses based solely on schoolwide student growth.

- Teachers in nontested grades and subjects can tie their performance to a core-subject area's classroom-level student growth.

One of the longest running performance-based compensation programs, the Professional Compensation System for Teachers (ProComp) in Denver, Colorado, is an example of a comprehensive

hybrid program. Teachers are eligible for additional compensation based on the following:

- Increasing their knowledge and skills through National Board for Professional Teaching Standards certification, participating in professional development, and earning graduate degrees;

- Receiving a satisfactory evaluation based on the district's teacher evaluation that takes place every three years (collaborative incentives are also available based on a school performance framework);

- Increasing student achievement at both classroom and schoolwide performance levels;

- Teaching in hard-to-staff schools or hard-to-fill subject areas.

The Austin Independent School District implements the REACH program, which provides incentives to teachers based on a variety of measures, including the following:

- Student growth, which recognizes teachers and principals for growth at the individual and schoolwide levels.

  - For individual classrooms, the district calculates student growth using student learning objectives.

  - For all teachers, the district calculates student growth using schoolwide achievement on TAKS (Texas Assessment of Knowledge and Skills).

  - The district calculates professional growth, which includes professional development such as the National Board of Professional Teacher Standard's (NBPTS) Take One! and novice teacher mentoring and recruitment and retention at high-need schools.

## What Does the Research Say About Reward Structures?

This section examines the research base in two key areas. These areas provide essential insight into what makes certain types of reward structures more attractive to states and districts looking to implement performance-based compensation and what potential methods states and districts should avoid. The first part of this section focuses on the perceived strengths and weaknesses of different types of programs. The second section explores the research available on stakeholder perspectives.

### Strengths and Weaknesses of Different Types of Reward Structures

In many cases, the research base indicates that the strength of one type of reward structure is the weakness in the other type of reward structure.

Research conducted in fields other than education yields some interesting perspectives. In one study, American workers across industries expressed more interest in individual performance-based awards than in awards based on the performance of groups (Kuhn & Yockey, 2003; LeBlanc & Mulvey, 1998). Other studies from the general American workforce suggest that workers prefer group incentives (Cable & Judge, 1994; Gomez-Mejia & Balkin, 1989; Kirkman & Shapiro, 2000). In these studies, however, it is important to note that the researchers do not clearly indicate the reasons behind this preference (e.g., frustration with individualized pay models or lack of interest in individualized pay, or other reasons).

Studies focused solely on educators indicate the existence of a comparable pattern. In two studies, educators indicated that they preferred awards based

on their students' performance to awards based on schoolwide performance, citing their level of influence on the results. Specifically, they felt more comfortable receiving or not receiving awards based on their own work because they felt it accurately reflected their effort, instead of receiving awards based on the efforts of others (Bretz & Judge, 1994; Kuhn & Yockey, 2003).

In follow-up work, Milanowski (2007) conducted a set of surveys and focus groups that uncovered further evidence of this trend. In this case, candidates in teacher preparation programs expressed concern about the idea that their salaries may be in the hands of their colleagues. Specifically, they believed that it is more equitable to reward effective teachers based on their individual work because school-based performance awards might lead to less effective teachers obtaining awards with minimal effort (Milanowski, 2007). The literature often refers to this concept as the "free-rider" problem.

Not all researchers, however, agree that the "free-rider" issue is teachers' most significant concern. A common criticism of individualized incentive plans is that they might negatively affect school culture by encouraging teachers to engage in uncooperative, competitive behaviors, such as withholding information or assistance from their peers (Keys & Dee, 2005; Murnane & Cohen, 1986; Organisation for Economic Cooperation and Development, 2009). A recent evaluation report from the Texas Educator Excellence Grant program reported that a small percentage of teachers (18.5 percent) indicated a sense of enhanced competition between teachers. In contrast, more than 80 percent of teachers reported a sense of duty to cooperate and support their fellow teachers in the program (Springer et al., 2009b).

Other research indicates that there is not always a sense of competition in schools that base performance awards on individual teacher performance. In an evaluation of the Little Rock, Arkansas, Achievement Challenge Pilot Program (ACPP), a program based on individual teacher awards, teachers in the program shared their satisfaction with their salaries in comparison to other teachers. Further, teachers did not report any counterproductive competition and believed that the work environment became more positive after ACPP implementation (Barnett, Ritter, Winters, & Greene, 2007).

As previously noted, the research that favors group performance awards hinges on the belief that this structure encourages teacher collaboration, whereas individual awards foster teacher competition (Azordegan, Byrnett, Campbell, Greenman, & Coulter, 2005). A recent report exploring views of Generation Y teachers (teachers who are 30 and younger) and some older generations of teachers found that 56 percent of both groups either somewhat or strongly favored school-based performance awards (Coggshall, Ott, Behrstock, & Lasagna, 2009).

Another theory is that having one or two effective teachers in a school will increase motivation and raise the standards for all teachers in the school. In one study, based on schools in North Carolina that use a group reward structure, Jackson and Bruegmann (2009) examined the impact that effective teachers (defined by their classroom-level student achievement gains) had on all teachers in the school. The researchers found that in the schools studied, student achievement in mathematics and reading rose across all classrooms in the same grade as the effective teachers. The researchers call this the "spillover effect" and posit that reward structures based on individual performance might reduce this effect (Jackson & Bruegmann, 2009).

Jackson and Bruegmann (2009) also present some theories on what creates the spillover effect. For example, teachers who share duties outside the classroom may interact with their more effective peers, which can reduce the burden of the shared tasks, allowing teachers to spend more time on personal professional development. Further, the presence of effective teachers may motivate other teachers to research and experiment with new teaching methods and improve their own performance.

As detailed in previous paragraphs, the research that illuminates the varied strengths and weaknesses of different reward structures tends to focus on either individual or group rewards only. In some cases, the researchers suggest that hybrid programs, programs that include both individual and group aspects, may be an answer to the challenges faced by individual- or group-only programs (Azordegan, et al., 2005; Coggshall, Ott, Behrstock, & Lasagna, 2009). Other researchers (Barnett et al., 2007; Milanowksi, 2007; Springer et al., 2009b) do not specifically point to hybrid programs; instead they recommend that policymakers pay close attention to stakeholder attitudes toward the proposed reward structures, as that has a greater effect on the relative success or failure of a program.

**Stakeholder Attitudes and Opinions on Reward Structures**

In addition to considering the available research on the strengths and weaknesses of individual and group reward structures, decisionmakers need to consider what research tells us about stakeholder attitudes toward both the different reward structures specifically and toward compensation reform generally. Stakeholder engagement and support is essential to the success of any alternative compensation program. This subsection highlights research on key stakeholder groups, teachers and the teachers unions, and their attitudes toward performance-based compensation and different

program structures. In order for performance-based compensation programs to have a chance of success, decisionmakers in the school system must consider these factors.



*... the presence of effective teachers may motivate other teachers....*

Both of the national teachers unions, the American Federation of Teachers (AFT) and the National Education Association (NEA), have expressed tentative support for certain types of alternative compensation in recent years. The AFT offers a resolution on professional compensation for teachers. This resolution includes support for a new compensation system that provides professional pay that moves beyond the "'rigid hierarchy' of the traditional salary schedule" (AFT, 2010). Specific to reward structures, the AFT expresses support for compensation systems that reward teachers in groups or as an entire school. Further, it suggests that districts or states base awards on multiple measures of student outcomes. The AFT

also indicates that it is more likely to support a program that provides multiple ways for teachers to earn additional compensation beyond student outcomes (AFT, 2010). The NEA has indicated that it does not support pay based on student test scores (Flannery & Jehlen, 2008), meaning it is also wary of individual reward structures based on student test scores. In a more recent report, the NEA expresses hesitant support for some forms of alternative compensation, while cautioning against others (Little, 2009). Specifically, Little, in an NEA-published paper (2009), asserts that before alternative compensation can work, districts and states need to have valid, reliable, and standards-based evaluation systems in place.

In some school systems, district or state officials collaborate with, and thus receive the support of, their local teachers unions. One such program is the Minneapolis Alternative Teacher Professional Pay System. One of the main forces behind the design and implementation of the program came from conversations between the district and the local teachers union. The partnership between the Minneapolis Public Schools and the Minneapolis Federation of Teachers established a baseline trust between the teachers and the district, which strengthened compensation reform efforts and facilitated implementation (Potemski & Rowland, 2009).

Research has shown that teacher attitudes toward compensation reform appear to follow a pattern—certain teacher and program characteristics have a significant association with their likelihood to support compensation reform. Specifically, the values and beliefs held by teachers, along with their perceptions of the fairness of the school or district program, affect their attitudes toward performance-based compensation.

One common thread found in the research is the idea that teachers are more likely to support a reward structure that motivates changes in behavior and, more specifically, behaviors that teachers value (Heneman, Milanowski, & Kimball, 2007; Springer et al., 2009a, 2009b). For example, as noted previously, Milanowski (2007) discovered that some teachers believe that individual teachers who contribute more deserve higher pay—these programs value extra effort. Further, research on the Texas Educator Excellence Grant and GEEG programs by the National Center on Performance Incentives indicates that more than half of teachers believe that incentives can encourage them to work more effectively and should capture professional practices that they consider are deserving of a monetary incentive (Springer et al., 2009a, 2009b).

Teachers are also more likely to support reward structures that they consider fair (Heneman et al., 2007). For example, the authors suggest that programs should reflect "distributive fairness" in order to be successful (p. 6). Distributive fairness refers to the extent to which teachers perceive the form, amount, and formula of an alternative compensation program to be reasonable. As mentioned previously, Gen Y teachers supported alternative compensation programs in the study conducted by Coggshall et al. (2009); however, they expressed concern about implementing individual performance-based bonuses, regardless of the measures of performance. The researchers found that 70 percent of Gen Y teachers felt it is unfair to link teacher pay solely to factors beyond their control, such as student demonstration of knowledge (Coggshall et al., 2009). These teachers (as well as other teachers in older generations) suggest that the ability of students to do well on opportunities to demonstrate knowledge, such as standardized tests, is often mitigated by negative issues at home that they cannot help but be affected by while in school (e.g., family turmoil or poverty).

## Considerations for Deciding Which Program Structure to Implement

The second goal of this paper is to provide districts and states deliberating implementation of a performance-based compensation program with some considerations CECR believes are essential to the planning process. The first important consideration is school and district culture. This consideration ties to the previous discussion on the research about the extent to which certain program structures promote or devalue collaboration. Assessing school or district culture is an important step to deciding whether to focus a program on individual or group performance. In addition, data system infrastructure and the inclusion of teachers from untested grades and subjects, especially teachers of special populations, are two practical factors to contemplate when deciding on the structure of an alternative compensation program. Below is a review of these three important considerations to deciding on program structure.

### School and District Culture

Performance-based compensation is highly dependent on school-based factors. The stakeholders directly affected by the program include the teachers, students, and in some cases the principal in each school implementing the program. As previously discussed, the attitudes and opinions of stakeholders can vary greatly, depending in some cases on the type of reward structure chosen. Therefore, determining which reward structure is right for a school, district, or state relies on a comprehensive understanding of the views held by the local teachers, principals, unions, and to a lesser degree students and/or parents.

These attitudes and opinions combine to create a culture in the school or district. The process used to determine the reward structure for any alternative compensation program is crucial to the overall program design and implementation process. Program planners should study programs using all types of reward structures to learn more about strengths of specific structures as well as design flaws. Relevant stakeholders should be familiar with the body of research on alternative compensation program design and implementation. Finally, program implementers should pay close attention to the history of reform in the district, which will inform the sense of culture and thus may affect the type of program that the district should implement.

Researchers are still attempting to understand this aspect of school and district culture on performance-based compensation reforms. Early reports in the field indicate that there is not enough research to understand fully the effect of different reward structures on teacher collaboration versus teacher competition, which might affect attitudes and opinions toward the program type chosen (Podgursky & Springer, 2007). In one study on incentive pay in India, researchers examined the effect of both individual and schoolwide incentives on student achievement (Muralidharan & Sundararaman, 2006) and found that (1) teacher incentives had a small, positive effect on student achievement results, and (2) there was no difference between gains associated with individual teacher incentives and schoolwide teacher incentives. The study also noted that results from the study are specific solely to the evaluated program. Researchers agree that programs have different effects, based on location, implementation, and local context, i.e., school and district culture.

### Data System Infrastructure

In addition to looking at school culture, school systems must also consider whether or not their existing data systems can support their compensation reform plans. The ability of a state, district, or school

to implement performance-based compensation relies on its ability to maintain and use the data collected to make decisions about teacher effectiveness and compensation. It is imperative for states, districts, and schools to have a firm understanding of their data system infrastructure in order to make decisions about program design. For example, if a district is unable to link individual teachers and their students, then it will be impossible to design a program whereby it can reward teachers individually as measured by student performance.

In the past, district data systems had a fairly simple role: provide the district with necessary compliance data for the state and Federal governments. The information collected was fairly simple and captured data within school buildings and districts, such as budgetary information, human resources data, assessment, and student information systems. Performance incentive initiatives place additional stress on the system, as school systems require their student information systems and human resources databases to link to each other. Ideally, school systems also should link assessment data to student information systems and human resources data. However, in most cases, each of these databases is separate and insulated. Student information system databases keep track of student enrollment, their schedules and courses, attendance, and in some cases, disciplinary actions. Human resources systems generally house employee data, deliver payroll, and track assessment and accountability requirements at both the school and grade levels. This separation between the two data systems, often referred to as data silos, makes using the data as a determinant of teacher pay more difficult and time-consuming (Thorn, Glover, & Watson, 2009; Watson, Kraemer, & Thorn, 2009).

Districts implementing alternative compensation systems should carefully examine the capacity of their information systems to make payouts at the grade and classroom levels. During this process, districts should determine how well the information system is able to (1) provide information on data quality essentials; (2) correctly provide data points for students and teachers at the course, grade, and school levels; and (3) establish strong processes for making student-teacher linkages.

Districts implementing group performance awards face challenges in data quality and analytic infrastructure. In many cases, the same challenges that prevent data systems from making linkages at the classroom level also make it difficult to analyze data at the grade or subject levels. However, some districts report that they have built their information systems to provide more accurate information at the schoolwide performance level. In many cases, it is still difficult to provide accurate performance data at the grade or subject level (Watson et al., 2009). Group awards may have an advantage over individual awards if the group measurement consists of student growth at the school level.

Performance-based compensation systems, especially at the individual reward level, have specific technology requirements, particularly when using value-added measures or other student growth measures. Stakeholders must consider these needs during the planning and design phase of any alternative compensation system to ensure that the program is valid and reliable. Districts and states need to be able to make links between data sets in order to match teachers with students. This approach requires not only a data infrastructure but also an analytic infrastructure comprising individuals with specialized skills in information technology to manage and support the system.

Watson et al. (2009) has determined the following six dimensions as essential for schools, districts, and states to be able to determine which teachers should receive an award, based on what information.

- **"Accuracy** is the degree to which data reflect reality" and are correct. Common challenges in this area include ensuring correct linkages between student scheduling and human resources systems and "poorly designed computer interfaces that do not check the validity of data at entry" (p. 4).

- **"Validity** is the degree to which data measure an intended construct…. One common example in education is the student school of record. While most students do not change schools during an academic year, many do, especially in urban settings. Thus, the school at which students participate in testing may not be the school at which they received most of their instruction" (p. 4).

- **"Granularity** is the number of individuals (e.g., students), items (e.g., test questions), or period of time (e.g., semester versus monthly attendance)" over which staff aggregate the data (e.g., a student who spends time in multiple schools, whether the district has sampled school-student data frequently enough) (p. 5). It is particularly important to establish the amount of time and effort assigned to each student in the many potential arrangements of team teaching, pullouts, specialist teachers, student mobility, teacher mobility, and so on.

- **Interoperability** is the degree to which multiple data systems have the ability to share the same data, such as merging teacher certification data from human resources with teacher course assignment data from student information systems.

- **Relational** is the degree to which an information system can depict the complexities of schooling (e.g., alternative approaches to school scheduling such as team teaching).

- **Reducibility** is the degree to which stakeholders can categorize data (e.g., teaching a particular subject or grade).

In addition to data system infrastructure as a consideration for program design, districts and states must also consider how to measure performance for eligible program participants, especially those who teach in untested grades and subject areas.

### Measures of Teacher Performance: Teachers of Untested Grades and Subject Areas and Teachers of Special Populations

Programs that use individual performance plans generally base at least a percentage of the award on student performance as measured by standardized achievement tests. Because of this, states, districts, and schools must consider the challenges related to measuring performance of teachers in nontested grades and subjects and develop systems that use alternative measures of student performance to assess teacher effectiveness.

The most significant challenge in implementing individual performance awards is identifying which variables the school system should use to measure teacher effectiveness and determine awards. Current education policy encourages the use of "multiple measures" to determine teacher effectiveness. With the influx of Federal funds available to states under the *American Recovery and Reinvestment Act of 2009* and specific requirements for states to receive funding, some state legislatures have mandated that, along with valid and reliable formative teacher observations, school systems must include student achievement measures as a "significant factor" in

teacher evaluation. In some cases, states indicate that evaluation results based on student growth may help inform decisions about teachers' careers, including compensation. Some researchers also suggest that evaluation systems should use multiple measures, including student achievement measures (Bill & Melinda Gates Foundation, 2010; Blanton, Sindelar, & Correa, 2006; Little, 2009).

This new pressure exacerbates an ongoing discussion within the compensation reform community—how to measure student achievement and growth at the individual level for the teachers of untested grades and subject areas, or as CECR refers to them, the "Other 69 Percent" (Prince et al., 2009). This category can include teachers of noncore subjects, such as art, music, physical education, foreign languages; teachers of prekindergarten to grade 2 and high school; and teachers of special populations, including English Language Learners (ELLs) and students with disabilities. This subsection focuses on implementing reward structures for teachers of these special populations. Research has shown four main challenges to evaluating teachers of special populations:

- Validity and reliability
- Use of value-added measures
- Teacher responsibilities and student diversity
- Local policy

### Validity and Reliability

One challenge to measuring individual teacher performance for teachers of special populations is the lack of valid and reliable student content assessments. There are two main reasons for the lack of validity. First, ELLs may understand a concept, but their English proficiency level may prohibit them from understanding the question terminology. Students with intellectual and developmental disabilities may require the use of alternative assessments, and thus, value-added models may not include them for similar reasons. Establishing

reliable and valid methods to determine teacher effectiveness for special student populations is necessary before teachers can receive individual performance-based rewards (Holdheide, Goe, Croft, & Reschly, 2010). Due to this challenge, teachers of ELLs and students with disabilities usually are not included in individual performance-based compensation systems. In order to include teachers of ELLs and students with disabilities, states and districts must use alternative measures of student performance beyond content-based assessments, which can provide challenges, particularly financial ones, as it may be time consuming and expensive to create these alternative measures.

### Use of Value-Added Measures

In a recent report, Braun, Chudowsky, and Koenig (2010) determined that it is uncertain how much of a teacher's value-added score truly represents the teacher's contributions and how much reflects the differences between classrooms in areas such as curriculum, resources, school culture, and peer effects. In essence, the researchers claim that much of a value-added score actually represents classroom effects versus teacher effects. For example, some studies have shown that small student samples commonly associated with special education and ELL caseloads produce results that are statistically less reliable (Amrein-Beardsley, 2008; Blanton et al., 2006; Feng & Sass, 2009; Koretz, 2008). In addition, even if calculated, value-added scores based on unreliable content assessment scores present a significant validity challenge. This creates difficulties when programs based on individual rewards rely on value-added or other student growth measures. Some subject areas and grade levels do not have assessments for use in value-added or student growth models. While not impossible to include these teachers in performance-based compensation systems, each school and district must consider these challenges and provide solutions while planning program design and structure.

## Teacher Responsibilities and English Language Learners

Another challenge to crediting this group of teachers for student learning, whether using a value-added system or other student growth measure, is determining how to measure the individual contributions of teachers who act as coteachers or consultants within a general education classroom (Holdheide et al., 2010). Schools and districts often have a wide variety of instructional options for ELL education. Teachers can take on different roles such as providing pull-out services, participating in full-day bilingual programs, receiving support from a bilingual paraprofessional, or serving as coteachers in a dual language program (Potemski, 2009). In addition, the student population in ELL services is extremely diverse. Teachers of these students are responsible for students with a range of special challenges and languages, which can make instruction and teacher effects difficult to normalize. When deciding on an alternative compensation program structure, it is imperative to take stock of all eligible participants. If teachers who have coteaching or consulting roles in the classroom are eligible to receive an award then program developers will have to decide how to measure these teachers' performance.

## Local Policy on Teaching Special Education Students

Like ELL teachers, special education teachers can be difficult to include in a performance-based compensation model. Like their ELL colleagues, special education teachers teach a very diverse group of students in an equally diverse number of possible settings. A recent study by the National Comprehensive Center for Teacher Quality, with assistance from the Council for Exceptional Children, surveyed state and district directors of special education and uncovered some significant challenges to measuring teacher effectiveness for teachers of special populations (Holdheide et al.,

2010). The survey revealed that a large percentage of districts (81.4 percent) have contracts that prohibit modifications to the evaluation process for special education teachers. To further complicate matters, a similar percentage of respondents indicated that special education teachers need to have knowledge, skills, and expertise that general education teachers do not. However, half of the survey respondents also expressed the opinion that teachers of special populations should not be evaluated using the same process as general education teachers (Holdheide et al., 2010). Measuring performance for teachers of students with special needs is often a complicated and challenging process; if states or districts structure an alternative compensation program in such a way as to allow these teachers to be eligible for performance-based rewards, then it will be important for program designers to be familiar with local policy issues in this area.



*When deciding on an alternative compensation program structure, it is imperative to take stock of all eligible participants.*

Frequently, states and districts respond to the challenges of providing individual performance-based awards to teachers of special populations by offering options for group-based performance awards. The main challenge to this approach is that there is not enough evidence to suggest the degree to which noncore teachers contribute to gains in other content areas (Koedel, 2007). In response, states and districts often base group awards on schoolwide performance awards in order to fairly reward all teachers in the school. While this often provides a temporary solution to providing performance-based compensation to all teachers, it is also unlikely that teachers in nontested grades and subject areas feel that this accurately reflects their contributions to student learning. In recent years, states and districts such as Austin, Texas, and Denver, Colorado, have experimented with the development of student learning objectives (SLO), which some educators believe allow for a more accurate measure of special educators' contributions to student learning. SLOs are statements that describe what a student should be able to do after participating in a learning activity. As research emerges from these programs, states and districts may have more options for structuring their performance-based compensation programs.

## Conclusion

As experimentation in compensation reform gains momentum throughout the country, states and districts must make decisions about the design and implementation of performance-based compensation programs. This paper focuses on the unit of accountability—individual versus group-based performance awards—and presents a research base to help states and districts determine which reward structure is right for their program. The most important considerations are the school culture, which includes stakeholder attitudes and opinions; the status of the data system and infrastructure; and issues related to the measurement of student

performance, which includes the use of value-added measures and teachers in untested grades and subjects:

- The attitudes of stakeholders toward teacher pay and alternative compensation programs can directly affect the likelihood of successful program implementation and sustainability. Because of this, the process used to determine the reward structure for any alternative compensation program is crucial to the overall program design and implementation process. Stakeholder groups should study programs using all types of reward structures to learn more about strengths of specific reward structures as well as design flaws. Program implementers also should be familiar with the body of research on alternative compensation program design and implementation. As noted in previous sections, researchers find that teachers are more likely to support compensation efforts when they feel that the program is fair and equitable, which is why so many districts opt for hybrid structures.

- Each district implementing an alternative compensation system, especially one that includes performance-based compensation, should carefully examine the capacity of its information system(s) to make payouts at the *grade* and *classroom* levels.

- Individual reward structures tend to ignore teachers in nontested grades and subject areas due to unreliable measures of student performance. Based on the interplay between perceived strengths and weaknesses of individual versus group reward structures, researchers often note that hybrid structures provide an opportunity to reward all teachers regardless of whether they teach nontested grades and subjects or special populations, while still supporting recognition for the most effective teachers.

# References

American Federation of Teachers. (2010). *Professional compensation for teachers.* Washington, DC: Author. Retrieved September 28, 2010, from http://www.aft.org/about/resolution_detail.cfm?articleid=205

Amrein-Beardsley, A. (2008). Methodological concerns about the education value-added assessment system. Educational Researcher, 37(2), 65–75.

Azordegan, J., Byrnett, P., Campbell, K., Greenman, J., & Coulter, T. (2005). *Diversifying teacher compensation.* Denver, CO: Education Commission of the States. Retrieved September 28, 2010, from http://www.ecs.org/clearinghouse/65/83/6583.pdf

Barnett, J. H., Ritter, G. W., Winters, M. A., & Greene, J. P. (2007, January). *Evaluation of year one of the achievement challenge pilot project in the Little Rock Public School District.* Fayetteville, AR: University of Arkansas, Department of Education Reform. Retrieved September 28, 2010, from http://www.uark.edu/ua/der/Research/merit_pay/LRSD_Report.pdf

Bill & Melinda Gates Foundation. (2010). *Empowering effective teachers: Strategies for implementing reforms.* Seattle, WA: Author.

Blanton, L. P., Sindelar, P. T., & Correa, V. I. (2006). Models and measures of beginning teacher quality. *The Journal of Special Education, 40*(2), 115–127.

Braun, H., Chudowsky, N., & Koenig, J. (2010). *Getting value out of value-added: Report of a workshop.* Washington, DC: National Research Council Committee on Value-Added Methodology for Instructional Improvement, Program Evaluation, and Accountability.

Bretz, R. D., Jr., & Judge, T. A. (1994). The role of human resource systems in job applicant decision processes. *Journal of Management, 20*(3), 531–551.

Cable, D. M., & Judge, T. A. (1994). Pay preferences and job search decisions: A person-organization fit perspective. *Personnel Psychology, 47*(2), 317–348.

Coggshall, J. G., Ott, A., Behrstock, E., & Lasagna, M. (2009). *Retaining teacher talent: The view from Generation Y.* Naperville, IL: Learning Point Associates and New York: Public Agenda. Retrieved September 28, 2010, from http://www.learningpt.org/expertise/educatorquality/genY/Gen%20Y%20report.pdf

Feng, L., & Sass, T. R. (2009). *Special education teacher quality and student achievement.* Report prepared for the U.S. Department of Education. Teacher Quality Research.

Flannery, M. E., & Jehlen, A. (2008, March). Where is your pay plan heading? *NEA Today.* Retrieved September 28, 2010, from http://www.nea.org/home/ns/4221.htm

Gomez-Mejia, L. R., & Balkin, D. B. (1989). Effectiveness of individual and aggregate compensation strategies. *Industrial Relations, 28*(3), 431–445.

Heneman, H. G., III, Milanowski, A., & Kimball, S. (2007). *Teacher performance pay: Synthesis of plans, research, and guidelines for practice* (CPRE Policy Brief RB-46). Philadelphia: Consortium for Policy Research in Education. Retrieved September 28, 2010, from http://www.cpre.org/images/stories/cpre_pdfs/RB46.pdf

Heyburn, S., Lewis, J., & Ritter, G. (2010). *Compensation reform and design preferences of Teacher Incentive Fund grantees.* Nashville, TN: Vanderbilt University, National Center on Performance Incentives. Retrieved September 28, 2010, from http://www.performanceincentives.org/data/files/news/PapersNews/2010_Heyburn_etAl_TeacherIncentiveGrantees.pdf

# References

Holdheide, L., Goe, L., Croft, A., & Reschly, D. J. (2010). *Challenges in evaluating special education teachers and English language learner specialists* (TQ Research & Policy Brief). Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved September 28, 2010, from *http://www.tqsource.org/publications/July2010Brief.pdf*

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

Keys, B. J., & Dee, T. S. (2005). Dollars and sense: What a Tennessee experiment tells us about merit pay. *Education Next, 5*(1), 60–67.

Kirkman, B. L., & Shapiro, D. L. (2000). Understanding why team members won't share: An examination of factors related to employee receptivity to team-based rewards. *Small Group Research, 31*(2), 175–209.

Koedel, C. (2007). *Teacher quality and educational production in secondary school* (Working Paper 2007-02). Nashville, TN: Vanderbilt University, National Center on Performance Incentives. Retrieved September 28, 2010, from *http://www.performanceincentives.org/_data/files/news/PapersNews/Koedel_2007a_Revised.pdf*

Koretz, D. (2008, Fall). A measured approach: Value-added models are a promising improvement but no one measure can evaluate teacher performance. *American Educator*, 18–39.

Kuhn, K. M., & Yockey, M. D. (2003). Variable pay as a risky choice: Determinants of the relative attractiveness of incentive plans. *Organizational Behavior and Human Decision Processes, 90*(2), 323–341.

LeBlanc, P. V., & Mulvey, P. W. (1998). How American workers see the rewards of work. *Compensation & Benefits Review, 30*(1), 24–28.

Little, O. (2009). *Teacher evaluation systems: The window for opportunity and reform.* Washington, DC: National Education Association.

Max, J. (2007). *The evolution of performance pay in Florida*. Washington, DC: U.S. Department of Education, Office of Elementary and Secondary Education, Center for Educator Compensation Reform.

Milanowski, A. T. (2007). Performance pay system preferences of students preparing to be teachers. *Education Finance and Policy, 2*(2), 111–132.

Muralidharan, K., & Sundararaman, V. (2006). *Teacher incentives in developing countries: Experimental evidence from India.* Cambridge, MA: Harvard University, Department of Economics.

Murnane, R., & Cohen, D. (1986). Merit pay and the evaluation problem: Why most merit pay plans fail and few survive. *Harvard Educational Review, 56*, 1–17.

Organisation for Economic Cooperation and Development. (2009). Design components of incentive pay programmes in the education sector. *OECD Education & Skills, 37*(24), 58–94.

Podgursky, M. J., & Springer, M. G. (2007). Teacher performance pay: A review. *Journal of Policy Analysis and Management, 26*(4), 909–949.

# References

Potemski, A. (2009). *Teaching English language learners: A complex system* (Policy-to-Practice Brief). Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved September 28, 2010, from *http://www.tqsource.org/pdfs/TQ_Policy-to-PracticeBriefELL.pdf*

Potemski, A., & Rowland, C. (2009). *Pay reform in Minneapolis Public Schools: Multiple approaches to alternative compensation* (Case Summary). Washington, DC: Center for Educator Compensation Reform. Retrieved September 28, 2010, from *http://cecr.ed.gov/pdfs/summaries/MinneapolisCaseSummary.pdf*

Prince, C., Schuermann, P. J., Guthrie, J. W., Witham, P. J., Milanowski, A. T., & Thorn, CA. (2009). *The other 69 percent: Fairly rewarding the performance of teachers of nontested subjects and grades.* Guide to implementation: Resources for applied practice. Washington, DC: Center for Educator Compensation Reform. Retrieved September 28, 2010, from *http://cecr.ed.gov/guides/other69Percent.pdf*

Springer, M. G., Lewis, J. L., Podgursky, M. J., Ehlert, M. W., Taylor, L. L., Lopez, O. S., et al. (2009a). *Governors Educator Excellence Grant (GEEG) program: Year three evaluation report.* Nashville, TN: Vanderbilt University, National Center on Performance Incentives. Retrieved September 28, 2010, from *http://www.performanceincentives.org/_data/files/news/BooksNews/GEEG_Year_Three_Report1.pdf*

Springer, M. G., Lewis, J. L., Podgursky, M. J., Ehlert, M. W., Gronberg, T. J., Hamilton, L. S., et al. (2009b). *Texas Educator Excellence Grant (TEEG) program: Year three evaluation report.* Nashville, TN: Vanderbilt University, National Center on Performance Incentives. Retrieved September 28, 2010, from *http://www.performanceincentives.org/_data/files/news/BooksNews/TEEG_Year_Three_Report1.pdf*

Thorn, C., Glover, R., & Watson, J. (2009). *Information technology considerations, Guide to implementation: Resources for applied practice.* Washington, DC: Center for Educator Compensation Reform. Retrieved September 28, 2010, from *http://cecr.ed.gov/pdfs/guide/itConsiderations.pdf*

Watson, J., Kraemer, S. B., & Thorn, C. (2009). *Data quality essentials, Guide to implementation: Resources for applied practice.* Washington, DC: Center for Educator Compensation Reform. Retrieved September 28, 2010, from *http://cecr.ed.gov/pdfs/guide/dataQuality.pdf*

# Performance-Based Compensation Structures:

Considerations for Individual, Group, and Hybrid Programs

CECR

Center for
Educator Compensation
Reform