2009

Merit and Performance-Based Pay

Nathan A. Bramble
St. John Fisher College

How has open access to Fisher Digital Publications benefited you?
Follow this and additional works at: http://fisherpub.sjfc.edu/education_ETD_masters
Part of the Education Commons

Recommended Citation

Please note that the Recommended Citation provides general citation information and may not be appropriate for your discipline. To receive help in creating a citation based on your discipline, please visit http://libguides.sjfc.edu/citations.

This document is posted at http://fisherpub.sjfc.edu/education_ETD_masters/64 and is brought to you for free and open access by Fisher Digital Publications at St. John Fisher College. For more information, please contact fisherpub@sjfc.edu.
Merit and Performance-Based Pay

Abstract
Merit-based pay provides teachers with monetary bonuses for reaching identified student achievement goals. Performance-based pay provides other types of bonuses to teachers for reaching similar student achievements. This paper will discuss both the positive and negative outcomes for providing these incentives to teachers. School districts in several states have implemented various aspects of merit and performance-based pay with varying results.

Document Type
Thesis

Degree Name
MS in Special Education

Department
Education

Subject Categories
Education
Merit and Performance-Based Pay

Nathan A. Bramble

St. John Fisher College
Abstract

Merit-based pay provides teachers with monetary bonuses for reaching identified student achievement goals. Performance-based pay provides other types of bonuses to teachers for reaching similar student achievements. This paper will discuss both the positive and negative outcomes for providing these incentives to teachers. School districts in several states have implemented various aspects of merit and performance-based pay with varying results.
Merit-based pay is defined as the practice of rewarding teachers with monetary bonuses for the accomplishment of target goals and/or basing a percentage of the teachers’ salary on student achievement. A merit-based pay program can be implemented in numerous capacities varying in percentages of salary and size of the bonus. Performance-based pay is defined as accepting a lower base pay in return for the opportunity to earn bonuses predicated upon predetermined goals. In contrast to the merit-pay system, performance-based compensation models seek to promote corporation and partnership among teachers. Rather than comparing teachers’ performance against each other, teachers are evaluated against a set of criteria determined by the school or school district.

The purpose of a merit and performance-based pay system is to enhance student achievement by rewarding teachers for the accomplishment of preset goals or benchmarks. Merit or performance-based pay can be subdivided into three general categories: individual, grade level teams and school districts. Individual incentives reward a single teacher for their performance, grade level teams promote cooperation and the participants equally share the rewards or bonuses, and rewarding entire school districts recognizes and reinforces the fraternal, interdependent learning culture.

For the purpose of this paper, the terms merit-based pay and performance-based pay will be interchangeable.

Positive Performance-based pay can be viewed as an innovative compensation strategy that has the potential to increase student learning through the use of effective teaching
Merit and Performance-Based Pay

practices. Odden and Carolyn (2000) make the argument that performance-based pay will determine the overall quality of the teaching professionals. "When coupled with other initiatives - e.g., signing bonuses, greater overall pay, improved working conditions, - higher beginning salaries are even more powerful. Education will not be able to recruit its share of quality individuals unless it pays competitive beginning wages." Positive attributes of a merit-based pay program include school districts ability to "...attract and retain more highly skilled teachers, and be a more efficient use of the educational dollar." Milanowski (2006). In his paper for the Wisconsin Center for Education Research, Milanowski (2006) elaborates that financial resources need to be used to retain high performers and districts with less impressive facilities, and the ability to match competitive salaries could strengthen district ability to retain effective teachers. Effective teachers are identified and rewarded. Teachers who are less successful can seek effective teaching practices from colleges, raising the overall standard of quality. Teachers who are unable to attain the new standard could be pushed out of the profession. "Besley and Machin (2006) found that the existence of a performance premium associated with being a good head teacher resulted in poor head teachers leaving their jobs rather than accepting lower salaries."

In their article in the Peabody Journal of Education, Podgursky and Springer (2007) cite a study by Lazear that concludes "A performance pay system will tend to attract and retain individuals who are particularly good at being incentivized and repel those who are not." In addition to attracting and retaining effective teachers, performance based pay creates a climate that encourages professional development. "Vaughn (school district) has anecdotal evidence that its pay program is attracting high-
quality younger teachers because of its pay-for-performance orientation and teachers' ability to earn a higher salary faster” (Odden & Kelly, 2000).

“The introduction of performance related pay can also motivate employees to pursue professional development opportunities that previously offered little in the way of additional benefits for the individual. Productivity is therefore likely to improve both in the short run, because employees are working harder, and also in the longer run, as staff professional development generates further gains in productivity” (Prentice, 2007).

Speaking about a case study outside the realm of education, Lazear states “…although the incentive system raised the productivity of the typical worker employed, it also tends to raise the overall quality of the workforce” (Podgursky & Springer, 2007). A similar opinion is expressed by Lawler (2006) in a speech to The Albert Shanker Institute. “You attract to some degree, to the degree that you are at or above market. You lose people if you are below market.” Lawler continues “If you give everybody an increase, the poorer performers often get above the market and are retained because they essentially have no options that equal what they are making where they are.”

**Negative**

The problem of insulating teachers’ performance from their compensation is compounded as teachers accumulate seniority enhancing the incentive to remain in the profession. Podgursky and Springer (2007) state “...studies of teacher turnover consistently find that high-ability teachers are more likely to leave teaching than teachers of lower-ability.” In response to this startling information, it is reasonable to assume that high-ability teachers were lured away from education. This is not the case, according to
...a recent provocative study by Hoxby and Leigh (2004) which found evidence that the migration of high-ability women out of teaching from 1960 to the present was primarily the result of the “push” of teacher pay compression – which took away relatively higher earnings opportunities for teachers - as opposed to the pull of greater nonteaching opportunities” (Podgursky & Springer, 2007). The collision of degrading forces resulting from the traditional seniority pay scale have obscured standards of accountability and have created a compounding incentive for low performing teachers to stay in education.

There are unintended negative outcomes which become potential concerns in regards to individual teachers and the cumulative effects on the learning culture within schools. “...teachers rated other outcomes quite low in desirability. These undesirable outcomes included public criticism for not meeting goals, loss of professional pride for not meeting goals, risk to job security, intervention, putting in more hours, less freedom to teach things unrelated to goals, and more pressure and job stress (Kelley, Odden, Milanowski, & Henenman III, 2000). The negative feelings become especially troublesome if they eclipse positive or enthusiastic sentiments. “Putting in more hours and more pressure and stress were seen as more likely than any of the positive outcomes” (Kelley et al., 2000).

A significant obstacle in the merit-pay debate is teachers’ attitude of acceptance. “...teachers may regard the pay increases they receive near the end of their careers as a recompense for relatively low initial salaries: being “underpaid” early in their career is offset by higher pay later” (Milanowski, 2006). In his examination of students preparing to be teachers Milanowski (2006) was repeatedly confronted with a positive outlook concerning merit-based pay. He proposes “It may be that new teachers – those who have
not yet experienced the subjectivity of evaluation and the instability of programs and funding, and who have not become accustomed to the traditional pay schedule – might be more accepting of performance-based pay.” It would appear that the optimism of perspective teachers far exceeds that of the teachers who have experienced a performance-based pay plan. “Just the issue of merit pay, which Milton mentioned, produced a book about six years ago that found 3,000 research studies evaluating the impact of merit pay systems in work organizations. And I can report that the results were overwhelming negative. The systems were not shown to have positive impacts, on balance, in virtually any setting” (Lawler, 2006).

Measurement

If the foundation of teacher compensation is to be based on the merits of the teacher, a primary consideration is how do you quantify merit? Multiple avenues of quantification have been proposed, each with their own advantages and disadvantages. The most popular assessment measures of teacher performance include: students’ scores on state tests, principal observation and evaluations, an independent evaluation agency, teacher content knowledge assessments, master teacher observation and evaluations, and benchmarks and target goals.

In his critique of rewards systems, Lawler (2006) stresses the importance of “…objective, creditable measures – and here’s the key point – inclusive of the major behaviors that you want someone to demonstrate. Many reward systems end up being dysfunctional because they pick an obvious easy-to-measure behavior…”
The process for evaluating teachers through observation is often criticized as being too subjective. State tests are viewed as a limited expression of a teachers' contribution in the classroom, and test scores are considered vulnerable to uncontrollable variables not related to education, such as illness, family conflicts, inadequate sleep, and countless others. "It was difficult to create a reliable process for identifying effective teachers, measuring the value-added to a student by an individual teacher, eliminating unprofessional preferential treatment from the evaluation processes, and standardizing assessment systems in schools" (Podgursky & Springer, 2007). A valid assessment system must comprehensively encompass teaching standards that produce student achievement.

"...a recent careful review by researchers at Rand concluded that studies "provide evidence that teachers have discernable, differential effects on student achievement, and that these effects appear to persist into the future," the size of these effects are difficult to determine, and many claims of big impacts on student achievement are exaggerated (Haskins & Loeb, 2007).

Milanowski (2004) based his assessment system on seventeen performance standards grouped into four domains. "Teachers were evaluated in six classroom observations and a portfolio prepared by the teacher. Four of the evaluations were conducted by a teacher evaluator, and two were conducted by principals and assistant principals. Conclusions from the study of more than 3000 teachers from more than seventy inter-city schools in Cincinnati indicated" "...teacher evaluation scores had a moderate degree of criterion-related validity." Another conclusion of the Milanowski (2004) study stated "...teacher evaluation scores may be useful as representations of
teaching practices that affect student learning.” One disclaimer of the Milanowski (2004) study that has a profound impact on the interpretation of data stresses “…it is important to recognize that very high correlations between teacher evaluation scores and student achievement measures are unlikely to be found for reasons including error in measuring teacher performance, error in measuring student performance, lack of alignment between the curriculum taught by teachers and the student tests, and the role of student motivation and related characteristics in producing learning.” If effective teacher practices cannot be conclusively linked to student achievement due to unreliable measurements of teacher and student performance, then how is it possible to determine merit or base a compensation model on the same error prone measurements. It is imperative that teachers have full confidence that the value of their input can be reliably measured and quantified.

“Often called value-added modeling, the general idea of the method is to use complex statistical techniques and repeated testing of students to measure changes in students to measure changes in student performance…But problems arise in using student test scores to identify effective teachers. The use of test scores gives teachers an incentive to manipulate the system by teaching test-taking skills, focusing more on some students than others, undermining the performance of other teachers, or simply cheating” (Haskins & Loeb, 2007).

The reliability of standards based instruction is further complicated by a claim from Ramirez (2001) in his article “How Merit-Pay Undermines Education”. He states “In theory, a first-year chemistry teacher fresh out of college is not as effective with high
school students as is a 10-year veteran with a master’s degree in science … critics counter that empirical evidence doesn’t support such logic. They assert that a teacher’s training is irrelevant and that experience counts for nothing.” Ramirez (2001) goes on to question “Unless all inputs are equalized for all teachers and administrators, how can policy makers judge the value of the outcome?” Podgursky and Springer (2007) also conclude there is a lack of continuity between estimated teacher effects and measured teacher characteristics. Lawler (2006) speaks to the unreliability of accurately quantifying a teacher’s individual role related to student achievement. “…”objective” metrics often lead to strange results that lead so many organizations to come back to a judgment of the individual made by a supervisor rather than a metric that is based on a test.” The idea that a subjective assessment trumps an objective quantification in relation to student learning is highlighted by Podgursky and Springer (2007). “A number of value-added studies find that principal evaluations are a reliable guide to identifying high- and low-performing teachers as measured by student achievement gains.” A study conducted by Sanders and Horn reinforces the reliability of principal evaluations. “There is a very strong correlation between teacher effects as determined by data and subjective evaluations by supervisors” (Podgursky & Springer, 2007).

One of the first large-scale studies, published by William Sanders and June Rivers of the University of Tennessee in 1996, was based on test score results in mathematics for students who were followed from grades three through five. Teachers for each grade were divided into five groups of equal size based on the improvement they produced in their students’ math scores. Students who had teachers in the top fifth of teacher
effectiveness for each of the three years scored about 50 percentile points better than students who had teachers in the lowest fifth. Subsequent analyses showed that teachers in the top fifth produced improvement among all students, regardless of their original scores or ethnic group. Teacher quality is the single most important feature of the schools that drives student achievement.” (Haskins & Loeb, 2007).

Concern about the integrity and independence of subjective evaluations must be considered. In a summative report examining the perspective of head teachers in England and Wales the “...possibility of unreliable decisions where the personal relationship between team leader and teacher is either very good or dislocated by previous tensions...” (Haynes, Wragg, Wragg, & Chamberlin, 2003). Lawler (2006) raises questions about the so called objective assessment of tests. “…most of the pay-for-performance systems highlight test scores, not teacher instruction, and test scores are much more subject to corruption than instructional practices.”

A clear system that accurately measures effective teaching practices has yet to emerge in the debate of merit-based pay. A schism also exists in the effort to establish a cause and effect relationship between effective teacher practices and student achievement. Teacher evaluations should be extensive to increase the validity of an accurate assessment. Given the lack of reliability from a myopic view of assessment, multiple avenues of assessment need to be employed with consideration to short-term and long-term growth.
“...research fails to establish any relationship between graduate degrees and student learning unless the degree is in the field in which the teacher specializes, and even here the evidence is quite weak. Similarly, there is no correlation between student learning and workshops, institutes, and study groups for teachers. Hill recommends that professional development be reformed to comply with three guideposts: courses or workshops should be at least several days duration; they should focus on subject-matter instruction; and they should have goals and curriculum materials used by the school system in which the teacher works” (Haskins & Loeb, 2007).

South Carolina School Incentive Reward Program (SIRP) is the longest running state-sponsored, group-based performance plan in the nation. The typical winning school receives between $15,000 and $20,000 (Kelly & Odden, 1995). A school gain index (SGI) is calculated for each school. Awards are based on three criteria, student achievement, teacher attendance, and student attendance. Of these three criteria, student achievement gain is the most important measure. Overall, schools have shown improvement in student performance on standardized exams. However, student and teacher attendance have not seen marked improvement. Schools in the lowest socio-economic status bands have seen the greatest improvements in student achievement.

Utilizing the Tennessee Value-Added Assessment System (TVAAS) as the measuring stick for students’ improvement and teachers’ productivity allows for a more “level playing field” for teachers who have low-achieving students. Tennessee Value-Added Assessment System focuses on students; improvement, and allows teachers that help previously underperforming students make significant learning gains during the year to receive credit for their accomplishments (Holland & Soifer, 2004). Rewards for high
TVAAS scores include a $5,000 bonus for individual teachers and the potential of a $2,000 bonus for every teacher in the school if the school receives a high overall TVAAS score. In addition to salary bonuses, other incentives provided to teachers by the CEA include loans toward the purchase of a house in a neighborhood near a low-performing school, free legal services, and free tuition toward a master's degree in urban education (Holland & Soifer, 2006). These incentives have helped reduce teacher recruitment and retention problems and improve student achievement at the nine schools utilizing this program (Plucker et al., 2005).

The following tables show the mean desirability responses and response distribution and the results for other measures of pay increase system preference. (Milanowski, 2006).

<p>| Performance Pay System Preferences of Students Preparing to Be Teachers Table 1 |
| Mean Desirability Responses and Response Distribution |</p>
<table>
<thead>
<tr>
<th>Pay increase based on:</th>
<th>Mean (std. dev.)</th>
<th>-4 or -3 highly undesirable</th>
<th>-2 or -1</th>
<th>0 neutral</th>
<th>+1 or +2</th>
<th>+3 or +4 highly desirable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual performance</td>
<td>1.7 (2.1)</td>
<td>5%</td>
<td>12%</td>
<td>4%</td>
<td>32%</td>
<td>46%</td>
</tr>
<tr>
<td>Developing knowledge and skills</td>
<td>1.8 (1.8)</td>
<td>3%</td>
<td>9%</td>
<td>10%</td>
<td>38%</td>
<td>40%</td>
</tr>
<tr>
<td>School performance</td>
<td>0.3 (2.3)</td>
<td>14%</td>
<td>25%</td>
<td>9%</td>
<td>32%</td>
<td>19%</td>
</tr>
<tr>
<td>Not based on performance</td>
<td>-0.4 (2.2)</td>
<td>23%</td>
<td>23%</td>
<td>18%</td>
<td>26%</td>
<td>10%</td>
</tr>
</tbody>
</table>
Performance Pay System Preferences of Students Preparing to Be Teachers Table 2

| Pay system                              | Rating of attractiveness vs. pay not based on performance | Percentage ranking of most desirable | Mean rank  
|-----------------------------------------|----------------------------------------------------------|-------------------------------------|------------
| Pay for individual performance          | 0.9 (1.0)                                                | 49%                                 | 1.8        
| Pay for knowledge and skill development | 0.8 (0.9)                                                | 30%                                 | 2.2        
| Pay for group performance               | -0.0 (1.1)                                               | 9%                                  | 2.9        
| Pay not based on performance            |                                                          | 12%                                 | 3.2        

Motivation

A key premise of the merit based pay system is the idea that teachers can be motivated by the opportunity to maximize their earning potential. Specific goals or teacher practices can be highlighted or emphasized with the benefit of increased focus creating a clear “line of sight” (Lawler, 2006) of desired objectives. Questions concerning the motivational potential of capital as it relates to teachers are an essential janus point. The answer to the question, “Are teachers motivated by money?” can be either a launch pad for the debate of merit-based pay or an impact barrier that terminates the discussion.

In a report summarizing the outcomes of school based performance awards from schools in Maryland, Kentucky, and North Carolina, researchers concluded “receiving a bonus was rated quite high in desirability, as were outcomes pertaining to recognition, satisfaction in meeting goals, seeing students learn and improve performance, and working with other teachers” (Kelley et al, 2000). Lawler (2006) fortifies the argument, “...performance-based reward that has some recognition value associated with it, a relatively small amount of money may be quite meaningful to people because it confirms
something about themselves; it carries a sense of status and a sense of recognition.” “It is much more likely that 9 or 10 percent or even 15 to 20 percent of pay needs to be at risk in order for it to be important enough for most people to get excited about and change their behavior” (Lawler, 2006). Other estimates that speak to the entire workforce are slightly more conservative. “Research in the private sector has found that in order to affect a worker’s motivation, annual bonuses need to be at least 5 to 8 percent of salary—about $2,000 for a typical teacher” (Odden & Kelley, 2000).

Teachers seem just as likely to be motivated by money as the next worker, but money alone can not maximize motivational potential. The synergetic interaction of monetary rewards and public recognition seem to be a packaged necessity in the institution of a successful performance-based program. The Heneman report cautioned against other relative and extraneous culminating factors that potentially impact their conclusions. “…the relationship between teacher motivation and school achievement we were likely to find might be weakened by many factors, including unreliability in the motivation and achievement measures and the impact of many factors other than teacher motivation on achievement” (Kelley et al., 2000).

Conclusion

The goal of a performance-based pay program is to increase effective teaching practices to elevate student achievement. The infusion of motivation into the profession of teaching is encouraging and exciting, but the unknown, unintended consequences of a merit-pay program need to be considered. The biggest obstacle performance-based pay faces is the subjectivity of assessments and observations and the validity of state and
teacher produced tests to measure student achievement. Quantifying acceptable and reliable data is a major road block on the path to merit pay. If an acceptable, reliable measurement of effective teacher practices and student achievement could be agreed upon, merit-pay could help improve student learning. Areas that need to be addressed are measurement and motivation, specifically targeting the quantification of student output as it relates to teacher input.
References


