

Paying Teachers According To Student Achievement: Questions Regarding Pay-For-Performance Models in Public Education

JAMES CAILLIER

Abstract: In an effort to correct for perceived deficiencies in the No Child Left Behind Act, value-added models were proposed as a way to find out how much students learned in schools and classrooms throughout the school year. What has garnered much controversy regarding the value-added model, however, is the attempt to link pay and tenure to performance. In this article, a theoretical framework is introduced that examines the likely success of using value-added assessments as a pay-for-performance tool. After applying the theory, the author suggests that caution should be used when utilizing value-added assessments to pay teachers for student learning.

Keywords: education, pay-for-performance, value-added, No Child Left Behind

Public sector performance management can be traced back to the 1980s, when local governments decided to implement private sector management strategies. These efforts, which were documented in a best-selling book by Osborne and Gaebler (1992), subsequently inspired President Clinton to use performance management concepts to overhaul the federal government (Hendrick 2000). Similar types of management reforms did not stop with Clinton. President Bush, upon taking office, announced the adoption of a results-oriented tool called the Performance Management Agenda (PMA) that seemed to build on what Clinton started (Milakovich and Gordon 2006). Although PMA primarily focused on federal agencies, the No

Child Left Behind Act of 2001 (NCLB) was, in part, a performance management tool that established federal standards that were passed down to state and local governments. Under NCLB, schools, local education authorities, and states were held accountable for student performance. This act literally shifted the responsibility for schools from local governments to the states and the federal government (Caillier 2007).

Since the passage of NCLB, state education agencies and legislators have worked to find a better way of holding schools accountable. The reason for these efforts is that NCLB's entire focus is on school-level, and not student-level, longitudinal data, and, as a result, school and teacher contributions toward individual learning, or the lack thereof, are masked (Misco 2008). Moreover, the data obtained for NCLB purposes do not assist in addressing issues around curriculum and practice (Doran and Fleischman 2005).

In an effort to correct these deficiencies, value-added models were proposed as a way to determine how much students learned in schools and classrooms throughout the school year. Despite the widespread appeal of these models, only one state—Tennessee—has embraced the idea in its entirety. Other states, like Ohio, are currently undertaking pilot studies to determine the effectiveness of this approach as a means of holding schools and teachers accountable (Misco 2008).

One issue of the value-added model that has attracted much controversy, however, is the attempt to link teacher pay and tenure to performance; that is, the dependence of teacher pay raises and tenure on

James Caillier, Ph.D., is an assistant professor of Public Administration in the School of Education and Human Services at The College at Brockport, State University of New York, Rochester, NY.

how well students in their classrooms perform. Thereby, the value-added model becomes a pay-for-performance tool, which is a mechanism in which employee pay is fixed to output. In this instance, teacher pay is tied to achievement scores (output).

Although similar merit pay reforms have been deemed unsuccessful in the past (Johnson 1984), some areas have moved toward this type of model, much to the dismay of teachers and their unions (Podgursky and Springer 2007). Most notably, New York City has begun a pilot project consisting of 2,500 teachers that, in the future, will connect personnel decisions (e.g., pay, tenure) to the value that teachers add to their students. In fact, some of the principals of participating New York City schools have not yet informed their teachers about these efforts (Medina 2008).

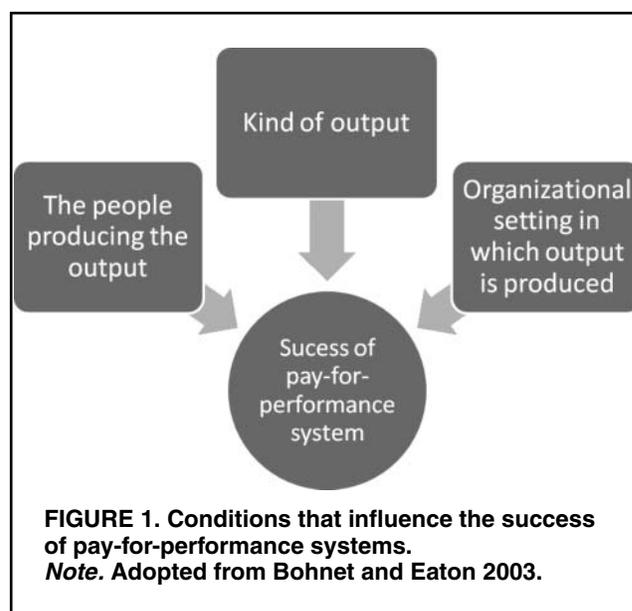
Although attempts to use the value-added model in personnel decisions will likely increase in the future, this article raises questions regarding perceptions of pay-for-performance models in educational settings. In this discussion, Bohnet and Eaton's (2003) pay-for-performance theoretical framework will be explained and related to public elementary and secondary education, despite the fact that the theory's initial application was to the evaluation of civil servants.

Three Necessary Conditions for Pay-For-Performance Success

Bohnet and Eaton (2003) provide a useful theoretical framework to illustrate the likelihood that a pay-for-performance system will be a success. Although *success* is a subjective term, it is defined here as motivating teachers to improve learning and adequately linking instruction to student learning, so that teachers can be financially rewarded accordingly. This model, which is depicted in figure 1, illustrates that three independent and distinct conditions determine the success of pay-for-performance systems. The first condition, the kind of output, specifies that outputs must be quantifiable and linked to a single employee and a single task. Second, the line staff, or those workers who produce the output, need to be motivated by pay. Last, the organizational setting in which the output is produced assumes that the workers are under one supervisor, and that they know exactly what they are supposed to produce. Each condition directly affects the success of the pay-for-performance system—meaning that deficiencies in these areas will hinder the system's effectiveness.

The Kind of Output Teachers Produce

When establishing a pay-for-performance system, the best results are produced when outputs are quantifiable and linked to a single assignment, and can be divided among employees (Bohnet and Eaton 2003). For the most part, teachers' primary responsibility is to in-



struct students. However, this instruction can encompass multiple subjects or assignments. For example, elementary teachers provide instruction in math, English, science, and social studies, with each of these classes being represented by student achievement. Therefore, elementary teachers have to be able to navigate from one subject to the next and to show improvement in at least two outputs (math and English), and possibly four in some states. Moreover, excellent teachers in one subject may not necessarily be excellent in another subject, and what is appropriate teaching for one student may not be appropriate for teaching another (Shulman 1986). Thus, it is quite possible for students to learn more in English and math than they do in science and social studies. It is possible that elementary teachers and schools will concentrate on certain subjects when their students are deficient in those areas. Middle and high school teachers, on the other hand, typically specialize in subjects. Nonetheless, elementary, middle, and high school teachers are more than instructors; they are also disciplinarians, counselors, and so forth, and they have to multitask when managing the classroom. Therefore, teachers perform many roles.

Achievement tests are aggregated based on achievement levels, percentages, or percentiles. Therefore, these tests are quantified and can be easily used to measure outputs. Achievement tests can also be linked back to the classroom teachers who deliver the instruction, provided that the appropriate student-level data is collected. Linking achievement results solely to a teacher's instruction, however, is an entirely different matter. Although research indicates that teachers make a substantial difference in learning (Rivkin, Hanushek, and Kain 2005; Sanders, Wright, and Horn 1997), studies also

show that class size (Nye, Hedges, and Konstantopoulos 2000; Smith, Molnar, and Zahorik 2003), peers (Coleman, et al. 1966), and socioeconomic status (Tajalli and Opheim 2005) may contribute as well. For example, affluent students may achieve gains when teachers are inadequate, and less affluent students may not realize targeted gains when teachers are excellent instructors. Just as the former results could be due to an educated parent who provides supplemental instruction, the latter could be caused by lack of support at home. Schools cannot do anything about a student's socioeconomic status, but they can seek to maximize or minimize factors such as class size that are under their control. There is also the possibility that classroom learning will result from pure chance. Additionally, to further complicate the value-added model, one specialized teacher can help another. For example, excellent teaching in English can lead to elevated math and science scores, because an improved command of the English language helps a student understand a text. For these reasons, research has found it difficult to quantify how much value a teacher adds to a student's learning (Podgursky and Springer 2007). Therefore, it is not enough for teachers to provide sound instruction; other factors need to be in place for the student to learn.

Additional questions surround the utility of achievement tests in measuring teacher effectiveness. For instance, value-added systems typically measure students' improvement from one year to the next via achievement tests. The idea is to employ sophisticated statistical techniques to determine the amount of value that a teacher adds. However, many factors, such as teaching to the test or the factors mentioned previously, can damage the linkage between these tests and effective teaching. For these reasons, some research questions the validity of value-added assessments in measuring teacher quality (Bracey 2004; Doran and Fleischman 2005; Martineau 2006).

As a result, other alternatives to measuring teacher effectiveness, including the benchmarks used by the National Board for Professional Teaching Standards (NBPTS), have been proposed to address these shortcomings. According to NBPTS, teachers should be able to "employ multiple measures of measuring student growth and understanding" and to "clearly explain student performance to parents" (National Board For Professional Teaching Standards 2002, 4). In essence, this process mandates that teachers must complete rigorous self-reflection exercises designed to demonstrate the effectiveness of their instruction. In at least one study, teachers with NBPTS certifications were shown to be more effective in raising student outcomes (Goldhaber and Anthony 2004). Therefore, schools should seek to use those models that best identify effective teachers.

The Teachers Producing the Output

Organizational literature is replete with approaches to motivation. These studies have found that private sector employees are motivated differently than those who work in public organizations (Crewson 1997; Perry 1996; Perry and Wise 1990). For example, civil servants were found to be more motivated by public service, or a desire to serve citizens, than their private sector counterparts. Studies have also documented that public sector employees are more motivated by non-monetary rewards (Borzaga and Tortia 2006; Ellickson 2002). More specifically, results from surveys have suggested that public school teachers are more motivated by work-related conditions than money (Hanushek and Rivken 2007). Furthermore, paperwork, student attitudes, lack of parental support, an unresponsive administration, and the low status of teachers were associated with teachers leaving the profession altogether (Tye and O'Brien 2002) or moving to a suburban school (Hanushek and Rivken 2007).

Conclusions like these limit the applicability of pay-for-performance mechanisms, which assume that money is the only motivator, and that it will inspire improved teaching. There are other reasons, however, why this model is supported: it makes it easier to provide professional development to deficient teachers, and to provide an alternative to NCLB. Nonetheless, if other factors serve to motivate teachers more, then monetary inducements alone will not maximize each teacher's instructional potential.

The Organizational Setting in which the Output is Produced

Leonard White advocated in 1929 that organizations should be structured so that every worker answers to a single individual at each level of the organization's structure (Milakovich and Gordon 2007). According to this theory, teachers should receive tasks, goals, and objectives from one individual, instead of several. For the most part, teachers do have a clear chain of command, and they know that they are under the authority of one principal. Specifically, a school's organizational chart is headed by the principal (the chief executive), and then followed by assistant principals (one of whom is the instructional leader), and finally teachers.

Even though schools have clear chains of command, teachers' clarity regarding goals and objectives may differ. This "can be attributed to multiple and/or changing leaders or managers with different objectives" (Bohnet and Eaton 2003, 18). Furthermore, additional policies addressing education are bound to be enacted by the government. This legislation will have an impact on organizational settings, because superintendents and principals will be mandated to implement different

policies. Therefore, school systems should constantly seek to clarify teachers' goals and objectives.

Conclusion

Value-added models have been recommended as a way to find out how much students have learned in classrooms throughout the school year. For that reason, this model will likely be adopted in the future. However, this system has several shortcomings that educators need to consider when linking the pay-for-performance of teachers to value-added assessments. First, many alternatives, besides value-added assessments, can be used to judge teacher effectiveness, and schools should employ those models that best identify good teaching. Moreover, additional school factors (class size, teacher quality, and so forth) can be used to increase learning outcomes, and schools should seek to adopt them. Second, factors other than pay (such as working conditions) serve to motivate teachers. Therefore, schools should use the motivational factors that will achieve the best results. Last, in changing organizational environments, teachers' understanding of goals and objectives is not a given. Thus, school systems should seek to clarify these expectations. In short, caution should be used when adopting the value-added model to pay-for-performance.

REFERENCES

- Bohnet, I., and S. C. Eaton. 2003. Does performance pay perform? Conditions for success in the public sector. In *For the people: Can we fix public service?*, ed. J. D. Donahue and J. S. Nye, 238–54. Washington, DC: Brookings Institutional Press.
- Borzaga, C., and E. Tortia. 2006. Worker motivations, job satisfaction, and loyalty in public and nonprofit social services. *Nonprofit and Voluntary Sector Quarterly* 35: 225–48.
- Bracey, G. W. 2004. Serious questions about the Tennessee value-added assessment system. *Phi Delta Kappan* 85 (9): 716–17.
- Caillier, J. G. 2007. The No Child Left Behind Act: Are states on target to make their goals? *Journal of Negro Education* 76 (4): 582–96.
- Coleman, J. S., E. Q. Campbell, C. J. Hobson, J. McPartland, A. M. Mood, F. D. Weinfeld, and R. L. York. 1966. *Equality of educational opportunity*. Washington, DC: U.S. Department of Health, Education, and Welfare.
- Crewson, P.E. 1997. Public-service motivation: Building empirical evidence of incidence and effect. *Journal of Public Administration Research and Theory* 7:499–518.
- Doran, H. C., and S. Fleischman. 2005. Challenges of value-added assessment. *Educational Leadership* 63 (3): 85–87.
- Ellickson, M. C. 2002. Determinants of job satisfaction of municipal government employees. *Public Personnel Management* 31:343–59.
- Goldhaber, D., and E. Anthony. 2004. Can teacher quality be effectively assessed? *Urban Institute Education Policy Center working paper*. Washington, DC: Urban Institute.
- Hanushek, E. A., and S. G. Rivkin. 2007. Pay, working conditions, and teacher quality. *Future of Children* 17 (1): 69–86.
- Hendrick, R. 2000. Comprehensive management and budgeting reform in local government. *Public Productivity & Management Review* 23 (3): 312–37.
- Johnson, S. M. 1984. Merit pay for teachers: A poor prescription for reform. *Harvard Educational Review* 54 (2): 175–85.
- Martineau, J. A. 2006. Distorting value added: The use of longitudinal vertically scaled student achievement data for growth-based, value-added accountability. *Journal of Educational and Behavioral Statistics* 31 (1): 35–62.
- Medina, J. 2008. New York measuring teachers by test scores. *New York Times*, January 21, A1.
- Milakovich, M., and G. J. Gordon. 2007. *Public administration in America*. 9th ed. Belmont, CA: Wadsworth.
- Misco, T. 2008. Was that a Result of my teaching? A brief exploration of value-added assessment. *The Clearing House* 82 (1): 11–14.
- National Board for Professional Teaching Standards. 2002. *What teachers should know and be able to do*. http://www.nbpts.org/UserFiles/File/what_teachers.pdf http://www.nbpts.org/UserFiles/File/what_teachers.pdf (accessed April 18, 2009).
- Nye, B. A., L. Hedges, and S. Konstantopoulos. 2000. Do the disadvantaged benefit more from small classes? Evidence from the Tennessee class size experiment. *American Journal of Education* 109:1–26.
- Osborne, D., and T. Gaebler. 1992. *Reinventing government: How the entrepreneurial spirit is transforming the public sector*. Reading, MA: Addison-Wesley.
- Perry, J. L. 1996. Measuring public service motivation: An assessment of construct reliability and validity. *Journal of Public Administration Research and Theory* 6:5–12.
- Perry, J. L., and L. R. Wise. 1990. The motivational bases of public service. *Public Administration Review* 50:367–73.
- Podgursky, M., and M. Springer. 2007. Credentials versus performance: Review of the teacher performance pay research. *Peabody Journal of Education* 82 (4): 551–73.
- Rivkin, S. G., E. Hanushek, and J. F. Kain. 2005. Teachers, schools, and academic achievement. *Econometrica* 73:417–58.
- Sanders, W., P. Wright, and S. Horn. 1997. Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Educational Research* 111:57–67.
- Shulman, L. S. 1986. Those who understand: Knowledge growth in teaching. *Educational Researcher* 15 (2): 4–14.
- Smith, P., A. Molnar, and J. Zahorik. 2003. Class-size reduction: A fresh look at the data. *Educational Leadership* 61:72–74.
- Tajalli, H., and C. Opheim. 2005. Strategies for closing the gap: Predicting student performance in economically disadvantaged schools. *Educational Research Quarterly* 28:44–54.
- Tye, B. B., and L. O'Brien. 2002. Why are experienced teachers leaving the profession? *Phi Delta Kappan* 84 (1): 24–32.

Copyright of Clearing House is the property of Taylor & Francis Ltd. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.