

What are achievement gains worth — to teachers?



School performance bonuses in New York City did not appear to inspire teachers to work harder or differently and had no effect on student outcomes.



By Julie A. Marsh and Daniel F. McCaffrey

Using financial incentives tied to performance has become a popular reform strategy in the education sector and beyond. Advocates of such incentives argue that they'll motivate educators to improve their practice and attract more teachers to the profession; detractors say they'll negatively affect morale and collegiality.

In the 2007-08 school year, the New York City Department of Education and the United Federation of Teachers introduced the Schoolwide Performance Bonus Program (SPBP). Unlike other pay-for-performance programs, groups of educators in schools were accountable for meeting program goals, not just individuals. The theory was that incentive pay would motivate educators to improve student achievement and that the chance to earn a bonus on the basis of school performance would enhance collaboration, leading to better outcomes.

Over two years, we surveyed teachers and staff in all participating schools and teachers in all of the eligible schools that were not selected for the program. We interviewed more than 130 individuals in 14 case study schools, interviewed program leaders, and analyzed administrative, individual, and school-level student achievement data. We issued our final report, *A Big Apple for Educators: New York City's Experiment with Schoolwide Performance Bonuses*, in July 2011 (Marsh et al., 2011).

We found that SPBP didn't improve schools or student outcomes. Student achievement didn't improve in schools randomly assigned to the program compared with control schools not assigned

JULIE A. MARSH (julie.marsh@usc.edu) is a visiting associate professor of education policy, University of Southern California, Los Angeles, and adjunct researcher at RAND Corporation. **DANIEL F. McCAFFREY** is senior statistician at RAND Corporation.



to it, and there were no differences in teachers' attitudes or practices, or the school climate. Moreover, as we were completing our study, two other studies on the effects of SPBP were released and reached a similar conclusion: The program did not yield positive effects on student achievement (Fryer, 2011; Goodman & Turner, 2011).

As a result of these studies, New York City abandoned its performance pay program in July.

Why didn't SPBP achieve its goals? And what does this imply for performance bonuses broadly? Simply stated, SPBP didn't create conditions necessary for performance-based incentive programs to change behaviors and didn't appear to motivate teachers to change their behaviors.

Overview of SPBP

Introduced in the 2007-08 school year, this voluntary program provided financial rewards to educators in high-needs elementary, middle, K-8, and high schools. Each school's union-represented employees — teachers, support staff, and counselors — voted whether to participate, and participating schools were eligible to receive school-level bonus awards equal to \$3,000 for each full-time, union-represented staff member to be distributed to staff if won by the school. Performance targets for awards were defined by the district's progress reports, the district's main accountability tool that measures student performance vis-à-vis standardized tests and comparisons with other schools, and the environment of all schools in the district. The program also required each participating school to establish a four-person compensation committee to determine how to distribute the bonus awards among staff.

In 2007-08, 427 high-needs schools were identified and about half were randomly selected to participate (called the "treatment" schools for the study's main analyses) and half not selected were called the "control" schools. There were 205 schools that participated in the first year, 198 schools in the second, and 196 in the final year. In the first year, 62% received bonuses totaling more than \$20 million; in the second, 84% of eligible schools earned more than \$30 million in awards. In the third year, due to the state raising its proficiency thresholds, only 13% of the schools earned bonuses totaling only \$4.2 million. The district suspended the program in January 2011 pending our study results and officially discontinued it in July upon release of our final report.

Key findings

The program did not improve student achievement at any grade level. The average mathematics and English language arts test scores in treatment

elementary, middle, and K-8 schools were lower than in control schools during all three years of the experiment. However, the differences were very small and statistically significant only for mathematics in the final year — and not significant when we controlled for testing effects from multiple years and subjects. Similarly, we found no overall effects on New York's Regents Exam scores for high school students in the first two years. Third-year data were not available for analyses. The program's effect did not differ among schools of different sizes, nor did it differ with the distribution plans for bonus awards among staff.

The program also did not affect school Progress Report scores.

Across all years and all categories of scores for the Progress Reports (environment, performance, progress, and additional credit), we found no statistically significant differences between scores of treatment and control schools. The lack of effects held true for elementary, middle, and high schools.

The program did not affect teacher-reported attitudes, perceptions, and behaviors.

Teachers in the treatment group and the control group showed virtually no statistically significant difference in reported instructional practices (e.g., using test data to guide instruction, aligning instruction with standards), effort (e.g., time spent on work outside regular school hours), participation in professional development, collaboration with other teachers or administrators, mobility (e.g., plans to return to their school next year), and attitudes (e.g., support for test-based rewards). Most teachers who received bonuses said the bonus did not affect their performance.

Most compensation committees developed nearly egalitarian award distribution plans, reflecting strong preferences among compensation committee members that staff members share bonuses equally.

Although administrators were significantly more in favor of differentiating bonus awards than their union counterparts, most committees developed essentially equal-share distribution plans, on average giving most staff members an award of about \$3,000. Although most plans included some small amount of differentiation for a handful of individuals — typically because individuals worked part-time or part of the school year — compensation committees were much less likely to judge individual performance when allocating bonus shares. Unequal disbursement at times resulted in resentment in schools, and some schools with highly differentiated allocation

plans one year adopted much more egalitarian plans the subsequent year.

Lack of positive effects

Two factors seem to have prevented the program from meeting its goals: The program did not develop conditions typically necessary for pay-for-performance programs to succeed, and bonuses had limited motivational effect.

Key supporting conditions were lacking.

Past research on pay-for-performance programs and expectancy theory (Chamberlin, Wragg, Haynes, & Wragg, 2002; Heneman & Milanowski, 1999; Kelley & Finnigan, 2003; Kelley, Odden, Milanowski, & Heneman, 2000; Vroom, 1964) suggests that to achieve the desired results, programs must strive to meet six conditions, and our study found mixed evidence of such conditions.

- 1. Understanding — Educators must have a high degree of understanding about how the program will function.** Although teachers said they knew about the program and generally supported it, more than a third didn't understand key elements, including targets their school needed to reach, bonus amounts, and how committees decided to distribute awards. Most teachers said they didn't know about distribution plans at the start of the year.
- 2. Expectancy — Educators must believe they're capable of doing work that will enable them to achieve the targets.** In the program's third year, staff seemed to understand what was required to earn a bonus, but they generally overestimated the likelihood their school would receive an award. For instance, although many compensation committee members acknowledged needing to improve performance to win a bonus in 2009-10, most felt certain their school would receive one — in the end, only 13% did.
- 3. Valence — Incentives must be sufficiently valuable to the potential recipients to inspire responses predicted by the theory of action.** Although most teachers expressed a strong desire to win an award, many winners reported that after taxes the amount seemed insignificant. In fact, almost half the teachers responding to the survey said the bonus wasn't large enough to motivate extra effort. Further, many case-study respondents saw the bonus as a reward for their usual efforts, not as an incentive for changing their behavior.

4. Buy-in — Educators must accept the program and its criteria for decision making. Most teachers and compensation committee members felt bonus criteria relied too heavily on test scores, indicating limited buy-in for program performance measures.

5. Fairness — Educators must perceive that teachers, staff, and schools will be treated fairly. More than one-half of teachers and compensation committee members felt the program was fair to participating schools, and compensation committee members did not report dissatisfaction about unfair distributions. However, some staff represented by the United Federation of Teachers (UFT) said they were unhappy with the composition of the compensation committee; more than one-half of teachers and UFT-represented members on the compensation committee wanted the UFT to have greater than 50% representation on the committee. Some staff also felt the Progress Report targets were too high.

6. Reasonable timeline — Educators must receive information with sufficient time to allow for responses predicted by the theory of change. Three-fourths or more of teachers and compensation committee members suggested that they should have been informed of the distribution plans at the start of the year.

Without a high degree of understanding, expectancy, valence, buy-in, perceived fairness, and a reasonable timeline, it is not surprising that SPBP was not a strong motivational and reform tool in all the schools.

The bonus had limited motivational power, particularly relative to other accountability incentives.

Assessing the true motivational value of the bonus was difficult, and our study yielded mixed findings on this topic. On the one hand, many teachers and staff said they wanted to earn the bonus and were taking it seriously. For example, 64% of teacher survey respondents said the possibility of a financial bonus motivated them moderately or more to work hard as a teacher. On the other hand many teachers and other staff said a possible bonus did not affect their practices. For instance, only 39% of compensation committee members and 15% of teachers reported that not receiving a bonus energized them to improve their practice the subsequent year.

These contradictions could have occurred because the bonus formula was too complex for teach-

ers to determine how much effort was needed to earn an award, as conjectured by another researcher who examined this program (Fryer, 2011) and supported by the large proportion of teachers confused about the Progress Report formula. Also, teachers may have lacked the capacity or resources to change, something that was conjectured by researchers who found no effects from a performance-pay experiment in Nashville (Springer et al., 2010).

Our data suggest a possible alternate interpretation: Financial rewards might not have motivated teachers to change because of how the rewards were viewed and because other factors are more salient to teachers. Principals and teachers in our case study schools consistently reported viewing the bonus as recognition for work their staffs were doing rather than a goal for which to strive. The bonus was frequently described as a “pat on the back” for a job well done. The principal at a school that won a bonus said, “It should be the icing on the cake, because we did the job. . . . It shouldn’t be the reason why we get the job done.” Similarly, most teachers and compensation committee members who actually received bonuses said that winning the bonus was a nice acknowledgement of their hard work but that it had not influenced their performance.

Other accountability pressures and intrinsic motivation also were often perceived to be more salient than the bonus. In surveys, teachers ranked intrinsic motivators — such as seeing students learn new skills and knowledge or seeing themselves improve as a teacher — much higher than financial bonuses on a list of potential motivators. Also, compared to receiving a financial bonus, a greater proportion of teachers said the possibility of receiving a high Progress Report grade (75%) and achieving or exceeding their school’s Adequate Yearly Progress target (77%) motivated them to work hard as a teacher. Thus, even if the bonus motivated staff to work hard or change their practice, they would probably have had similar motivation without it because of the high level of accountability pressure on all schools and their staffs. Consequently, the Schoolwide Performance Bonus Program might not be expected to change behavior or to influence student outcomes.

Policy implications

The results of our research add to a growing body of research from the United States that finds no effects on student achievement of narrow pay-for-performance policies that focus only on financial incentives without including other features such as targeted professional development or revised teacher evaluations (Glazerman & Seifullah, 2010; Springer et al., 2009; Springer et al., 2010). Our results yield several implications relevant to the broader set of

pay-for-performance policies that have received considerable attention in recent years:

- **Conditions must foster strong motivation.**

This study supports past research suggesting that certain conditions — timeline, understanding, valence, etc. — may be necessary to bolster the motivational effect of financial incentives. Bonus programs may need to communicate more effectively about their accountability apparatuses and consider the level of buy-in for and understanding of those measures. Staff may be unwilling to work toward a goal they don’t support, even if they find bonuses valuable. The system may have to engage staff in developing the performance metrics and be willing to revise them if they don’t receive general support or retain the initial buy-in.

Other related conditions that weren’t closely examined in this study but could be important are the capacity of educators to achieve the accountability targets and the organizational context in which educators operate. If educators feel that they don’t have the resources or that organizational or political barriers impede their ability to act in ways that will help them achieve their targets, then incentives will be of little use.

A growing body of research in the United States finds no effects on student achievement of narrow pay-for-performance policies that focus only on financial incentives without including other features such as targeted professional development or revised teacher evaluations.

- **Identifying factors that truly affect motivation is important.**

Motivation is the key to the theory of change in pay-for-performance programs. However, our evaluation presents a somewhat contradictory story. Teachers and other staff said the possible bonus motivated them to work hard and was desirable, but they also said the program did not inspire them to change their behavior. So, a desirable award may not be enough to change behavior.

This may be particularly true in an era of high-stakes and high-profile accountability. More teachers said that receiving a good Progress Report, rather than a bonus, was a motivator. The Progress Re-

port has no direct financial rewards, but it has high stakes in terms of public opinion and possible negative sanctions. Concerns about experiencing negative sanctions — again a factor faced by all schools — could be a greater motivator of true change than

Financial rewards might not have motivated teachers to change because of how the rewards were viewed and because other factors are more salient to teachers.

a bonus with unknown expected value. In the context of high-stakes accountability — some New York City schools were facing possible state takeover — student outcomes of both treatment and control teachers demonstrated notable improvements, but the bonus provided no additional effects on student achievement.

• Performance-based incentives may face challenges from the micropolitics of school-level implementation.

Although many major program elements were implemented smoothly across participating schools, some schools had difficulty deciding how to distribute bonuses among staff, and some unequal disbursements exacerbated political tensions in schools. The very idea of differentiating pay based on performance may challenge deeply ingrained norms of collaboration and egalitarianism in the profession.

• Pilot testing and evaluation are essential.

From the outset, the New York City Department of Education and teacher union leaders planned to pilot the program before scaling up. Resource constraints prevented the planned scale-up but we had three years to compare outcomes and practices across participating and comparable high-needs control schools. Beginning on a small scale and including randomized treatment and control groups for three years provided valuable information to inform future decisions about an essentially untested policy innovation. Those considering similar programs should allow for pilot testing and evaluation of the theory and assumptions underlying any new pay-for-performance program.

In the end, our study does not provide a definitive answer about whether performance bonuses are an effective school improvement strategy. But the

results do suggest the need to think carefully about designing incentive programs and creating conditions that foster the desired motivational effects, and to anticipate the likely challenges involved in achieving positive outcomes. ■

References

- Chamberlin, R., Wragg, T., Haynes, G., & Wragg, C. (2002). Performance-related pay and the teaching profession: A review of the literature. *Research Papers in Education, 17*, 31-49.
- Fryer, R.G. (2011). *Teacher incentives and student achievement: Evidence from New York City Public Schools*. Cambridge, MA: National Bureau of Economic Research. www.nber.org/papers/w16850
- Glazerman, S. & Seifullah, A. (2010). *An evaluation of the Teacher Advancement Program in Chicago: Year Two impact report*. Washington, DC: Mathematica Policy Research.
- Goodman, S. & Turner, L. (2011, Spring). Does whole-school performance pay improve student learning? Evidence from the New York City schools. *Education Next, 11* (2), 66-71.
- Heneman III, H.G. & Milanowski, A.T. (1999). Teachers' attitudes about teacher bonuses under school-based performance award programs. *Journal of Personnel Evaluation in Education, 12*, 327-341.
- Kelley, C.J. & Finnigan, K. (2003). The effects of organizational context on teacher expectancy. *Educational Administration Quarterly, 39*, 603-634.
- Kelley, C., Odden, A., Milanowski, A., & Heneman III, H. (2000). *The motivational effects of school-based performance (CPRE Policy Brief, RB-29)*. Philadelphia, PA: Consortium for Policy Research in Education.
- Marsh, J.A., Springer, M.G., McCaffrey, D.F., Yuan, K., Epstein, S., Koppich, J., Kalra, N., DiMartino, C., & Peng, A. (2011). *A big apple for educators: New York City's experiment with schoolwide performance bonuses*. Santa Monica, CA: RAND Corporation.
- Springer, M.G., Lewis, J.L., Podgursky, M.J., Ehlert, M.W., Gronberg, T.J., Hamilton, L.S., Jansen, D.W., Stecher, B.M., Taylor, L.L., Lopez, O.S., & Peng, A. (2009). *Texas Educator Excellence Grant (TEEG) Program: Year Three evaluation report*. Nashville, TN: National Center on Performance Incentives.
- Springer, M.G., Hamilton, L.S. McCaffrey, D.F., Ballou, D., Le,V., Pepper, M., Lockwood, J.R., & Stecher, B.M. (2010). *Teacher pay for performance: Experimental evidence from the Project on Incentives in Teaching*. Nashville, TN: National Center on Performance Incentives.
- Vroom, V.H. (1964). *Work and motivation*. New York, NY: Wiley.