



EDUCATION

CHILDREN AND FAMILIES

EDUCATION AND THE ARTS

ENERGY AND ENVIRONMENT

HEALTH AND HEALTH CARE

INFRASTRUCTURE AND
TRANSPORTATION

INTERNATIONAL AFFAIRS

LAW AND BUSINESS

NATIONAL SECURITY

POPULATION AND AGING

PUBLIC SAFETY

SCIENCE AND TECHNOLOGY

TERRORISM AND
HOMELAND SECURITY

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis.

This electronic document was made available from www.rand.org as a public service of the RAND Corporation.

Skip all front matter: [Jump to Page 1](#) ▼

Support RAND

[Purchase this document](#)

[Browse Reports & Bookstore](#)

[Make a charitable contribution](#)

For More Information

Visit RAND at www.rand.org

Explore [RAND Education](#)

View [document details](#)

Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Unauthorized posting of RAND electronic documents to a non-RAND website is prohibited. RAND electronic documents are protected under copyright law. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please see [RAND Permissions](#).

This product is part of the RAND Corporation monograph series. RAND monographs present major research findings that address the challenges facing the public and private sectors. All RAND monographs undergo rigorous peer review to ensure high standards for research quality and objectivity.

A Big Apple for Educators

New York City's Experiment with Schoolwide Performance Bonuses

FINAL EVALUATION REPORT

Julie A. Marsh, Matthew G. Springer, Daniel F. McCaffrey,
Kun Yuan, Scott Epstein, Julia Koppich, Nidhi Kalra,
Catherine DiMartino, Art (Xiao) Peng

Prepared for the Fund for Public Schools



RAND EDUCATION

VANDERBILT



Peabody College

This research was conducted by RAND Education, a unit of the RAND Corporation, and partners at Vanderbilt University and the National Center on Performance Incentives (NCPI). The research was sponsored by the New York City Fund for Public Schools and NCPI.

Library of Congress Cataloging-in-Publication Data

A big apple for educators : New York City's experiment with schoolwide performance bonuses : final evaluation report / Julie A. Marsh ... [et al].

p. cm.

ISBN 978-0-8330-5251-3 (pbk. : alk. paper)

1. Teachers—Salaries, etc.—New York (State)—New York. 2. Merit pay—New York (State)—New York. 3. School improvement programs—New York (State)—New York. 4. Academic achievement—New York (State)—New York. I. Marsh, Julie A.

LB2842.4.N7B54 2011

371.1009747—dc23

2011023662

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

RAND® is a registered trademark.

Cover illustration courtesy Getty Images

© Copyright 2011 RAND Corporation

Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Copies may not be duplicated for commercial purposes. Unauthorized posting of RAND documents to a non-RAND website is prohibited. RAND documents are protected under copyright law. For information on reprint and linking permissions, please visit the RAND permissions page (<http://www.rand.org/publications/permissions.html>).

Published 2011 by the RAND Corporation

1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138

1200 South Hayes Street, Arlington, VA 22202-5050

4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665

RAND URL: <http://www.rand.org>

To order RAND documents or to obtain additional information, contact

Distribution Services: Telephone: (310) 451-7002;

Fax: (310) 451-6915; Email: order@rand.org

Executive Summary

In the 2007–2008 school year, NYCDOE and UFT implemented the SPBP in New York City public schools. The RAND Corporation, a nonprofit public policy research institution—in a partnership with Vanderbilt University and NCPI, a U.S. Department of Education-funded national research and development center for state and local policy on educator compensation—conducted an independent evaluation of the SPBP for NYCDOE and the UFT, with funding from The Fund for Public Schools and NCPI. This two-year study took place over the period from February 2009 through March 2011 and addressed the following broad research questions:

1. How was the program being implemented from 2007 through 2010?
2. What were the intermediate outcomes of the program?
3. What were the effects of the program on student performance?

This report describes the results of our analyses for all three years of the program, from 2007–2008 through 2009–2010.

Overview of the Schoolwide Performance Bonus Program

Implemented for the first time in the 2007–2008 school year, the SPBP was a joint NYCDOE-UFT program to explore the utility of an alternative compensation plan as a means of improving student outcomes. The theory of action behind the SPBP was that an incentive pay system

would motivate educators to change their practices to ones better able to improve student achievement. The program's theory of action also hypothesized that the chance to earn a bonus on the basis of school performance could enhance collaboration and that winning bonuses could boost morale—both of which would lead to better outcomes. The theory of action that some leaders articulated held that, in addition to the overall motivational effect of the incentive pay system, rewarding staff differentiated awards on the basis their individual performances could provide extra incentives for change.

SPBP was a voluntary program implemented in high-needs elementary, middle, K–8, and high schools. For schools participating in SPBP, NYCDOE set annual performance targets based on its Progress Reports. The Progress Reports are the district's main accountability tool, measuring student performance and the school environment in all schools in the district. For elementary and middle schools, Progress Report scores are determined in large part by student growth on standardized tests and the school's performance standing relative to other schools, both their peers and all others citywide; scores for high schools are based on graduation and progress to graduation as measured by credit accumulation and performance on Regents Examinations. If a participating school met its annual performance target, as defined by the NYCDOE Progress Reports, the school could receive school-level bonus awards equal to \$3,000 per full-time UFT-represented staff member working at the school. The program required each participating school to establish a four-person compensation committee (CC) to determine how the bonus award would be distributed among staff (the CC distribution plan). The committee was made up of the principal, a member designated by the principal, and two UFT-represented staff members elected by staff.

In 2007–2008, Harvard Professor Roland Fryer, with input from the NYCDOE, identified 427 high-needs schools and randomly selected about one-half of these schools for the opportunity to participate in the SPBP in the first year. Fifty-five percent of school staff needed to agree to participate: 205 schools participated in the program in the first year; 198 schools remained in the second year, and 196 in the third and final year.

The program distributed more than \$50 million in bonuses over the three years of the study. In year 1 (2007–2008), 62 percent of participating schools won bonuses, resulting in more than \$20 million in awards. In year 2 (2008–2009), the numbers increased to 84 percent and more than \$30 million. In year 3 (2009–2010), the state raised the thresholds for proficiency, and correspondingly, the percentage of schools earning bonuses decreased significantly. That year, only 13 percent of participating schools earned bonuses, and just \$4.2 million were distributed. On announcing the year 3 bonus results in January 2011, NYCDOE simultaneously suspended the program, noting that it would base its decision about whether to resume the program on the results of this study.

Study Methods

For this evaluation, we collected a variety of qualitative and quantitative data and capitalized on SPBP's experimental design to evaluate program effects. They conducted surveys, site visits, interviews, and case studies to gather data about the implementation.

To understand the implementation and perceived effects of the program from the perspective of staff members closely involved with it, we surveyed all four members of the CCs in the 198 participating schools in spring 2009 (68-percent response rate across all respondents) and 196 participating schools in spring 2010 (72-percent response rate).

To gain an in-depth understanding of the SPBP and staff experiences with the program, we visited seven case-study SPBP schools in spring 2009 and seven additional schools in spring 2010. We selected the case-study schools to represent varying school and student characteristics (school level, student enrollment, borough, student achievement, and student demographics) and school experiences with SPBP (how the school's CC allocated the bonus, whether the school received a bonus, and whether the school voted to participate). Over the two years, we interviewed more than 130 individuals, including principals and assistant principals, teachers, instructional specialists, clinical and student support staff, secretaries, and paraprofessionals.

In addition, to understand SPBP's history, goals, and theory of action, we reviewed official program documents and interviewed NYCDOE and UFT representatives, funders, and other leaders.

To analyze student achievement and compare schools assigned to participate in SPBP with those not assigned to the program, we obtained NYCDOE administrative data for the baseline year (2006–2007) and all three years of implementation. These data included student- and school-level files with biographical, demographic, and student achievement information, along with publicly available data from school Progress Reports. NYCDOE also provided data on school participation, CC distribution plans, and bonus awards for each program year.

Finally, to test whether SPBP affected teachers' attitudes and behaviors (e.g., instruction, collaboration), in spring 2010, we surveyed a sample of teachers in all schools selected to participate in the SPBP and eligible schools that were not selected for participation (57-percent response rate across all respondents).¹

Key Findings

The Study Found No Effects on Student Achievement

Overall, SPBP Did Not Improve Student Achievement in Any Grade Level. Analyses of student achievement on the state's accountability tests found that the average mathematics and ELA test scores of students from elementary, middle, and K–8 schools randomly chosen for an invitation to participate in SPBP were lower than those of students from control schools during years 1, 2, and 3. The magnitudes of the estimates, however, are very small and statistically significant only for mathematics in year 3. The results are not significant when we controlled for testing effects from multiple years and subjects. These results were robust under various analytic approaches. Similarly, there were

¹ A PDF file containing the six appendixes to this document is available on the product page for this monograph (<http://www.rand.org/pubs/monographs/MG1114.html>). These appendixes offer an extensive collection of information on our surveys, bonus distribution, and student achievement data and analysis methods.

no overall effects on Regents Exams scores for high school students in years 1 and 2. We tested for but did not find differential program effects by school size and found no relationship between student achievement and the CC distribution plans for bonus awards among staff.

SPBP Had No Effects on School Progress Report Scores. Across all years and all the component scores for the Progress Reports (environment, performance, progress, and additional credit), we found no statistically significant differences between scores of SPBP treatment and control schools and between schools that participated in SPBP each year (regardless of random assignment) and other eligible schools. This lack of effects holds true for elementary, middle, and high schools.

The Implementation Had Mixed Success in Creating the Optimal Environment for SPBP

SPBP implementation depended on communications about the program to participating schools and staffs, the CC determination of award distributions, the determination of award winners, and the subsequent payout of those awards. Past research on pay-for-performance programs and expectancy theory (Vroom, 1964) suggests that, to achieve the desired program results, these activities must follow a reasonable time line and lead to a high degree of the following key attributes:

- *understanding* of the program, as shown by knowledge of criteria by which incentives are awarded and the amount of money at stake
- *expectancy*, as demonstrated by educators' beliefs that they are capable of doing things that will enable them to achieve the targets
- *valence*, a concept that refers to the belief that incentives are sufficiently valuable or substantial to inspire responses predicted by the theory of action
- *buy-in*, or acceptance of the program and its criteria
- perceived *fairness*.

The following subsections describe the rollout, communication, and the extent to which the implementation achieved the key attributes theoretically necessary for success.

Basic Procedures Were Enacted as Planned. The NYCDOE and the UFT generally implemented the SPBP as intended. At the start of each year, schools voted to participate and, later in the year, formed their CCs and submitted bonus distribution plans. Schools had freedom to compensate staff with bonuses as they desired, without interference from either NYCDOE or UFT. Each year, NYCDOE awarded bonuses according to the program guidelines. In the first two program years, district leaders announced the bonus awards within the first few months of the next school year. However, in year 3, announcements were delayed until nearly midway through the subsequent school year.

Communications Followed the SPBP Design, with Early Problems Being Corrected over Time. Both NYCDOE and UFT adhered to the communication plan for SPBP, which called for sharing joint NYCDOE-UFT written materials about the program with school staff. According to survey respondents, UFT representatives served as the main conduits for information about SPBP to UFT CC members and teachers. CC respondents reported early misunderstandings but that communications improved over time. Nevertheless, some interviews and teacher survey responses suggest that some misunderstanding of program components remained in year 3, when UFT and DOE reduced the emphasis on communicating about the program.

The CC Process Was Implemented Fairly and Smoothly, but Some Schools Had Difficulty with the Decisionmaking Process. Each school formed a four-person CC and generally followed SPBP guidelines on membership and procedures. Most CC members reported that the decisionmaking process was fair, collegial, and inclusive and that achieving consensus was easy; however, some survey and case-study respondents expressed concerns about the process. For example, 44 percent of teachers disagreed or strongly disagreed that teachers' preferences were taken into account when developing the distribution plan. Some CC members whom we interviewed also questioned whether the requirement that UFT-represented staff make up only one-half of the members truly guaranteed an even playing field and spoke

about power differentials that played out between administrators and UFT-represented staff in committee deliberations.

The Majority of CCs Developed Nearly Egalitarian Award Distribution Plans, Reflecting Strong Preferences Among CC Members for Equal Bonus Shares for All Staff. Although administrators were significantly more inclined than their UFT counterparts to favor differentiating the bonus awards so that different staff members would be eligible for different amounts, overall there was a strong preference for egalitarian plans among CC members. Further, almost two-thirds of teachers indicated a preference for distributing equal shares of the bonus to all school staff (yet almost the same proportion of teachers also reported that nonteachers should receive a smaller share than teachers). Not surprisingly, the majority of committees developed equal-share distribution plans in both years. The most common individual bonus amount awarded to staff in the plans was \$3,000 in all three years. Most staff within a school received the same award amount. In fact, in 2010, 82 percent of staff members were slated to receive the most common or modal award for their school, and even greater equality existed in the early years of the program. Inconsistent with the notion that larger schools might use differentiation in bonus payments to offset “free riders” (i.e., staff who shoulder less than their fair share of the work but still collect a full bonus), award equality increased with school size.

About 31 percent of schools reported using individual performance as at least one of the factors for determining awards. The remaining schools either did not differentiate or reported using only factors related to time or job title but not individual performance. Unequal disbursement at times resulted in resentment within the schools, and some schools with highly differentiated allocation plans one year adopted much more egalitarian plans the subsequent year.

The Few Schools That Determined Bonus Shares by Individual Performance Tended to Have More Differentiation Among Award Amounts but Did Not Differ from Other Schools in Student Achievement. Sixty schools reported including individual performance measures, such as staff absences and unsatisfactory staff evaluation ratings (“U-ratings”), as a factor (among many others) for determining individual staff awards. Compared to other schools that did

not report relying on individual performance measures in developing their distribution plans, these schools were more likely to award staff less than the modal awards and more likely to award certain staff members no money. According to the theory of action for SPBP some leaders espoused, greater disparity in awards would help to incentivize performance, have a motivational effect on individual behavior, and increase student achievement. However, even the schools that determined awards by individual performance generally remained cautious about deviating from egalitarian awards and slated 74 percent of staff, on average, for the modal award amount. Moreover, students in these schools—whether elementary, middle, or K–8—did not have higher achievement than those in other SPBP schools.

A Minority of CC Members and a Few Case-Study Respondents Reported Problems with the Distribution of Bonuses. The distribution process was reported to run fairly smoothly in the third year of the program, and only a minority of CC members reported problems, such as staff being left off the list of recipients or being upset with a perceived unfair distribution. None of these reported problems reached the level of official appeal in any year of the program. Participants in several case-study schools reported dissatisfaction among staff when decisions and criteria used to differentiate the bonus were viewed as opaque and subjective, or when the schools did not communicate the final distribution plans to staff members.

The Implementation of SPBP Had Mixed Results in Creating the Circumstances That Foster Success. We noted the following:

- *Understanding and Awareness.* Staff members reported being aware of the program and generally supportive of it. Most educators surveyed and interviewed greatly appreciated the financial reward and the recognition of their efforts. However, there were persistent misunderstandings about the Progress Reports and other program elements. According to surveys, more than one-third of teachers did not understand key aspects of the program, including the target their school needed to reach, the amount of money their school would receive if they met their target, the source of the funding, and how committees decide on distribu-

tion plans. In the case-study schools, many individuals conveyed misperceptions about various aspects of the programs, and many called for better communication.

- *Expectancy.* In the third year of the program, staff showed some accuracy in their assessment of the efforts required to earn a bonus but, overall, seemed to overestimate the likelihood their school would receive an award. For instance, although many CC members acknowledged needing to improve performance to win a bonus in 2009–2010, a large majority felt certain their school would receive one.
- *Valence.* The majority of teachers and CC members expressed a strong desire to win bonuses and found the financial bonuses motivating, but many winners reported that, after taxes, the bonus seemed insignificant. In fact, almost one-half of the teachers responding to the survey and some staff members in case-study schools indicated that the bonus was not large enough to motivate extra effort. Further, many case-study respondents reported viewing the bonus as a reward for their usual efforts, not as an incentive for changing their behavior.
- *Buy-in.* Buy-in for the program performance measure was limited: The majority of teachers and CC members felt the criteria relied too heavily on test scores.
- *Fairness.* More than one-half of teachers and CC members felt the program was fair to participating schools, and CC members did not report dissatisfaction about unfair distributions. However, some UFT-represented staff expressed dissatisfaction with the composition of the CC committee; more than one-half of teachers and UFT committee members wanted the UFT to have greater than 50-percent representation on the committee. Some staff also felt the Progress Report targets were too high.
- *Time lines.* Three-fourths or more of teachers and CC members suggested that they should have been informed of the distribution plans at the start of the year.

SPBP Did Not Produce the Intended Effects on Teachers' Reported Attitudes and Behaviors

SPBP Had no Effects on Teacher Reported Attitudes, Perceptions, and Behaviors. The survey found no differences in the practices and opinions of teachers in the SPBP group, as compared with those in the control group. On all but one measure related to instructional practice, effort, participation in professional development, mobility, and attitudes, the responses from the two groups of teachers were very similar, and there were no statistically significant differences among them.

Lack of Results Might Be Due to the Limited Motivational Power of the Bonus. The theory of action behind the SPBP is that the potential for a bonus would motivate change, but this study did not find such a change. The vast majority of teachers and CC members who received bonuses said that winning the bonus was a nice acknowledgement of their hard work but that it did not influence their performance. In addition, only 39 percent of CC members and 15 percent of teachers reported that not receiving a bonus energized them to improve their practice the subsequent year, and only a very small proportion of both groups actually reported that not receiving the bonus reduced their motivation.

The limited motivational effect of the bonus might have resulted from school staff viewing the award as a reward rather than an incentive: It made them feel appreciated for their hard work, but they claimed that they would have undertaken this work with or without the bonus. The size of the award might also have been a factor. Some stakeholders believed that differentiation among awards could enhance the motivational power of the bonuses. As discussed above, the differentiation especially related to performance was very limited, and it is possible this too contributed to the limited incentivizing effects of the bonus. Finally, other accountability pressures and intrinsic motivation were often perceived to be more salient than the bonus. For example, although 64 percent of teacher survey respondents said the possibility of earning a financial bonus motivated them moderately or to great extent to work hard as a teacher, an even greater percentage reported the same motivational value from the possibility of receiving a high Progress Report grade (without the financial incentive attached)

(75 percent) and achieving or exceeding their school's adequate yearly progress target (77 percent).

The Lack of Effect Was Also Echoed in Responses About Staff Members' Own Performance. The majority of teachers, CC members, and case-study respondents said that SPBP did not affect their own job performance or motivation.

It Is Less Clear Whether Lack of Program Effects on Collaboration or Morale Contributed to the Lack of Intermediate Outcomes. SPBP's theory of action hypothesized that the chance to earn a bonus on the basis of school performance could enhance collaboration and lead to better outcomes. It also suggests that winning bonuses could boost morale and lead to better outcomes. However, the data on collaboration and morale are mixed. More than 40 percent of participating teachers and CC members reported that teachers' willingness to collaborate and work together, teachers' focus on student learning, and staff morale changed for the better as a result of participation in SPBP. The perception that the program caused the increased willingness to collaborate, however, is not reflected in the comparison with schools not assigned to the program. The reported levels of staff collaboration differed very little between the SPBP and control schools, and the differences that did exist tended to show greater collaboration in the control schools (but the differences were not statistically significant).

What Explains the Lack of Change Under SPBP?

The findings from the student achievement analysis suggest that, in the three years of implementation, the SPBP did not achieve its ultimate goal of improved student performance. There are several plausible explanations for this result. First, one could argue that the program was too new to produce effects—especially in the first two years. However, if newness were the explanation, one would expect to see at least some positive effects by year 3, but this was not the case.

Second, the findings from the teacher and CC surveys and case-study interviews suggest that several factors important for pay-for-performance programs might not have been in place in all participating

schools. The absence of these factors might have weakened the motivational effects of the bonus. Circumstances that could have weakened SPBP's effects include misunderstandings about the program, uncertainty that teachers could sufficiently change their practices or act in ways to achieve targets, lack of buy-in for the bonus criteria, concerns about timing, and questions about procedural fairness.

A third possible explanation for the lack of positive student achievement effects is that there could be a flaw in the underlying theory of action for SPBP and school-based pay-for-performance programs more generally. As some have argued in the past, motivation alone does not improve schools. Even if the bonus inspired staff to improve practices or work together, it may not have the capacity or resources (e.g., school leadership, social capital, expertise, instructional materials, time) to bring about improvement.

Finally, the lack of observed results could be due to the low motivational value of the Progress Reports and the accountability incentives all schools face. Assessing the true motivational value of the bonus is a difficult task, and our study yielded mixed findings on this topic. While many teachers and other staff reported that the possibility of receiving a bonus did not affect their practices, many nonetheless expressed a desire to earn the bonus and indicated that they were taking it seriously. Nevertheless, many acknowledged that other accountability pressures and incentives (e.g., receiving a high Progress Report grade, achieving adequate yearly progress targets) held the same motivational value as the possibility of receiving a financial bonus.

Hence, while the bonus might have been another factor motivating SPBP staff to work hard or change their practices, they would probably have had similar motivation without it because of the high level of accountability pressure all schools and their staffs face. Consequently, SPBP might not be expected to change behavior or to influence student outcomes.

Recommendations

Overall, the evidence from this study does not support continuing SPBP. The program did not have the desired effects on student achievement or the necessary intermediate teacher outcomes. There is no consistent evidence that the program motivated educators to change practices or that continuing the program would improve outcomes. Moreover the program was costly, and costs fluctuated dramatically across years.

As was discussed earlier, some implementation problems may have weakened the effects of the bonus and the program overall, and NYCDOE and UFT leaders could consider the following actions to address those:

- Improve communication to participating schools about bonus criteria, how the program operates, potential funds available, and the committee process.
- Adjust the program time line to ensure that all staff know, at the start of the year, the distribution plan and the dollar amounts they can potentially earn.
- Ensure transparency of distribution criteria and plans.
- Provide technical assistance to CC members on the decisionmaking process for determining distribution plans.

However, we do not have evidence that implementing these adjustments will improve the effectiveness of the program significantly. Furthermore, given the consistent failure of recent experiments with similar bonus programs to find positive effects on student outcomes, we suspect a more comprehensive revision of the program and reevaluation of the theory of action is likely to be necessary to achieve the desired outcomes.

Finally, the sharp decrease in the percentage of schools receiving bonuses in year 3 (13 percent) compared to year 2 (84 percent) also suggests a broader lesson for policymakers generally. Although the study did not collect data on responses to the year 3 results, it is hard to imagine how such fluctuations could improve the already weak moti-

vational effect of the program. Prior to implementing a performance-based bonus program, it would behoove leaders to ensure greater stability of the accountability measures on which the bonuses rely.

Implications for Pay-for-Performance Policies

The results of this evaluation 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 other features, such as targeted professional development or revised teacher evaluations. In recent years, a few studies, released well after SPBP started, have demonstrated that several school- and individual-based performance incentive policies had no significant effect on student test results (e.g., Glazerman and Seifullah, 2010; Springer et al., 2009a; Springer et al., 2010a). The NYC SPBP provided one of the few examples of a school-based incentive policy and, like the individual programs, it resulted in few observable effects on students, staff, and schools. The implementation findings, nonetheless, provide insights into the theory of action underlying the policy and the mechanisms expected to yield positive student outcomes—such as improved motivation, enhanced morale, increased collaboration, and improved practice. The study also uncovered much about the process and potential challenges of assigning a small school committee the task of distributing a schoolwide bonus among staff.

Overall, these results yielded several implications relevant to the broader set of pay-for-performance policies that have gained considerable national attention and interest in recent years.

- *Conditions must foster strong motivation.* Our study findings support much of the theoretical literature and other empirical research suggesting that there may be a set of key conditions (e.g., a high degree of understanding, expectancy, valence, buy-in, perceived fairness, and a reasonable time line) needed to bolster the motivational effect of financial incentives. Several of these purported key system components were lacking in SPBP and were

identified by some educators as limiting the ability of program to change their behaviors.

- *It is important to identify the factors that truly affect motivation.* Motivation is the key to the theory of change in pay-for-performance programs. The failure of recent experiments to find effects (Springer et al. 2010a, b) has generally been attributed to the lack of motivational effects of such programs. However, the in-depth data from this evaluation present a somewhat contradictory story. Teachers and other staff reported that the possibility of a bonus was desirable and motivating, but they also reported not changing their behaviors in response to the program. It may be that a desirable award is not enough to actually change behavior. This may be particularly true in the context of high-stakes and high-profile accountability. More teachers reported that receiving a good Progress Report grade was a motivator than did teachers who reported the bonus motivated them. The Progress Report has no direct financial rewards, but it does have high stakes in terms of public opinion and possible negative sanctions. All schools face these accountability pressures, and these pressures may thus reduce the relative value of the additional incentive of the bonus to the point that it does not matter or that it seems merely to substitute for other motivational factors. It may also be that concerns about negative sanctions—again a possibility all schools face—is a greater motivator of true change than a bonus with unknown expected value.² Finally, as noted earlier, the limited motivational effects of the bonus reported by teachers and other school staff might not be enough to explain why SPBP failed to change teacher behaviors. Such factors as lack of capacity to change or to identify alternative teaching practices might have prevented change, even if the limited motivation of the SPBP bonus induced teachers to try to improve.
- *Performance-based incentives may face challenges from the micropolitics of school-level implementation.* This evaluation highlighted the

² The experience of the POINT experiment in Nashville, Tennessee, is consistent with this conjecture (Springer et al., 2010a).

underlying political tensions inherent in implementing a bonus system. Although many major program elements were implemented smoothly across participating schools (e.g., the formation of CCs), some schools found it difficult to decide how to distribute bonuses among staff (e.g., power differentials played out between administrators and UFT-represented staff during deliberations). The majority of CCs distributed bonuses nearly equally among all staff, and some unequal disbursements exacerbated political tensions within schools. Those seeking to enact similar programs should recognize that the very idea of differentiating pay based on performance will likely challenge deeply ingrained school cultures and norms of collaboration and egalitarianism.

- *Pilot testing and evaluation are essential.* From the outset, NYCDOE and UFT leaders intended to implement the SPBP on a pilot basis. Resource constraints prevented the planned scale-up within the initial group of schools identified as eligible for this pilot program, providing evaluators with three years to compare outcomes and practices across a set of participating and comparable high-needs control schools. Implementing the program on a small scale and including randomized treatment and control groups for three years (perhaps not intended at the outset) provided valuable information to inform future decisions about an essentially untested policy innovation. Leaders created the opportunity to gather comprehensive data on the implementation and outcomes to assess the merits of continuing and further scaling up this policy. The pilot period and evaluation also provided leaders an opportunity to reexamine the theory of action and assumptions underlying the SPBP.