



# 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](http://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](http://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.

# Improving School Leadership Through Support, Evaluation, and Incentives

---

## The Pittsburgh Principal Incentive Program

*Laura S. Hamilton • John Engberg • Elizabeth D. Steiner  
Catherine Awsumb Nelson • Kun Yuan*

Sponsored by the Pittsburgh Public Schools

The research in this report was produced within RAND Education, a unit of the RAND Corporation. The research was sponsored by the Pittsburgh Public Schools.

#### Library of Congress Cataloging-in-Publication Data

Improving school leadership through support, evaluation, and incentives : the Pittsburgh principal incentive program / Laura S. Hamilton ... [et al.].

p. cm.

Includes bibliographical references.

ISBN 978-0-8330-7617-5 (pbk. : alk. paper)

1. School improvement programs—Pennsylvania—Pittsburgh—Case studies.
  2. Educational leadership—Pennsylvania—Pittsburgh—Case studies.
  3. Educational change—Pennsylvania—Pittsburgh—Case studies.
  4. School principals—Pennsylvania—Pittsburgh—Case studies.
- I. Hamilton, Laura S.

LB2822.83.P43167 2012

371.2'07—dc23

2012021169

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 photograph courtesy iStock*

© Copyright 2012 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 2012 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](mailto:order@rand.org)

## Summary

---

In 2007, the Pittsburgh Public Schools (PPS) received funding from the U.S. Department of Education's Teacher Incentive Fund (TIF) program to implement the Pittsburgh Urban Leadership System for Excellence (PULSE), a set of reforms designed to improve the quality of school leadership throughout the district. A major component of PULSE is the Pittsburgh Principal Incentive Program (PPIP), a system of support, performance-based evaluation, and compensation. The new compensation system has two major components: (1) an annual opportunity for a permanent salary increase of up to \$2,000 based primarily on principals' performance on a rubric that is administered by assistant superintendents (who supervise principals) and that measures practices in several areas and (2) an annual bonus of up to \$10,000 based primarily on student achievement growth. The district also offered bonuses to principals who took positions in high-need schools. PPIP provided principals with several forms of support, such as professional development focused on improving leadership, feedback and coaching from assistant superintendents, and participation in Directed Professional Growth (DPG) projects that allowed principals to choose an area in which to work to improve their own skills. The RAND Corporation served as the evaluator of PPIP and examined implementation and outcomes from school years 2007–2008 through 2010–2011. Although the district is likely to continue implementing much of what constitutes PPIP, this report focuses only on the period during which PPIP was being funded by the TIF grant. The evaluation addressed the following broad questions:

- What is the district's theory of action regarding how PPIP is expected to promote improved student outcomes?
- How were the PPIP capacity-building interventions implemented, and how have principals responded to them?
- To what extent have principals' skills and practices changed over the course of PPIP?
- What conditions have changed at the school and classroom levels over the course of PPIP?

- How did principals perform on the rubric and bonus measures, and how was performance related to principal mobility?
- How did student achievement change during the course of PPIP, and how did racial and socioeconomic gaps change?

PPIP is one manifestation of the national focus on improving school leadership as a means of promoting effective teaching and student achievement. PPIP was implemented in a reform-oriented district in which it was one of many ambitious initiatives in place. The most salient addition to the slate of reforms during PPIP's implementation was a district-wide program to improve student achievement and student preparation for postsecondary goals by focusing on improving teacher effectiveness. In 2008–2009, the district developed the Research-Based, Inclusive System of Evaluation (RISE), a new system for observing and evaluating teacher practice and for guiding teacher professional growth. In 2009, based in part on the RISE work, the district received substantial funding from the Bill and Melinda Gates Foundation to implement the Empowering Effective Teachers (EET) program, which incorporated RISE. Many aspects of the EET work were consistent with PPIP. For example, the rubric used to evaluate principals encouraged principals to spend more time in classrooms observing and coaching teachers; the RISE process not only provided the framework for this task but also prompted changes to the principal rubric and to the supports and professional development provided to principals as part of PPIP.

As this discussion illustrates, the designs of EET and PPIP appear to be well aligned, thus providing a context that both supports PPIP implementation and complicates interpretation of the effects of PPIP. In particular, given the context of the district's multiple, overlapping reforms, it is impossible to disentangle any effects of PPIP from the effects of other reforms. Nonetheless, the information presented in this report should contribute to an improved understanding of how reforms like PPIP are implemented, how principals and other school staff respond to these reforms, and what outcomes might be expected to accompany their implementation.

## Data and Methods

The evaluation combined qualitative and quantitative data-collection and analysis approaches in an effort to provide a comprehensive understanding of the program theory, implementation, and outcomes. To understand the theory of action guiding the district's design of PPIP and to document the reform's implementation over time, we conducted focus groups and individual interviews with key district staff in all years of the evaluation. We also collected and reviewed extensive documentation, including meeting minutes; documentation of program components, such as the rubric; informa-

tion about principal professional development; and email exchanges with district staff throughout the four years of the evaluation.

To gain an in-depth understanding of principals' responses to PPIP, changes in principals' practices and skills, and changes in conditions at the school and classroom levels, we conducted interviews and surveys with principals and other school staff. We surveyed principals, online or in-person, during each year of the evaluation. We also surveyed school-level curriculum coaches in years 1, 3, and 4 of the evaluation and teachers in year 4. In addition, we conducted interviews with principals and other school staff in each year of the evaluation. The primary purpose of the interviews was to supplement the surveys by obtaining responses with richer detail and greater depth. To assess principal performance on the rubric and bonus measures, their responses to the high-need school incentive, and changes in student achievement, as well as gaps in achievement, we analyzed administrative data provided by PPS.

## Key Findings

### **What Is the District's Theory of Action Regarding How the Pittsburgh Principal Incentive Program Is Expected to Promote Improved Student Outcomes?**

PPIP relied on a combination of capacity-building interventions that included professional development in various forms, evaluation and feedback provided to principals by their supervisors, and financial incentives. District staff viewed the support and feedback interventions as having more potential impact on principal performance than the financial incentives, and principals themselves were more likely to attribute changes in their leadership to support and feedback than to incentives. District leaders posited that the interventions would affect principals by improving principal knowledge and skills and by influencing principals' practices, specifically by increasing time spent on instructional leadership activities. The district also expected PPIP to work by eventually improving the quality of candidates who apply for positions in the district while encouraging less effective principals to leave the district. These changes were then expected to promote learning-focused school environments, improved instruction, and higher levels of student achievement. Our data indicated that the multiple reforms taking place in PPS were perceived as working well together and providing a coherent set of policies focused on raising student achievement throughout the district.

### **How Were the Pittsburgh Principal Incentive Program Capacity-Building Interventions Implemented, and How Have Principals Responded to Them?**

Most principals reported that the PPIP-supported professional development contributed to their professional growth, and, in the final year of the evaluation, principals gave particularly high ratings to learning walks with their internal instructional leadership teams and participation in their own DPG projects. The district provided profes-

sional development through the Leadership Academy to promote principals' instructional leadership, and majorities of principals agreed that the Leadership Academy helped them improve their skills across multiple leadership domains, most strongly in the areas of monitoring teachers' instruction and providing feedback to teachers. In addition, nearly two-thirds of principals reported that one-on-one coaching from their assistant superintendents made a moderate or large contribution to their professional growth.

Principals viewed the rubric that was used to evaluate their practices as a useful tool for thinking about their strengths and weaknesses, and we found evidence that principals were increasingly accepting of the idea that unfavorable school conditions should not be considered an excuse for poor performance on this measure. PPS refined the measure over time by reducing the number of standards on the rubric and increasing its focus on principals' roles as instructional leaders and managers of human capital. Principals appreciated these changes and reported that the new rubric standards were well-aligned with the work they were doing to support and evaluate teachers. At the same time, majorities of principals expressed concerns about fairness throughout the course of the evaluation, including a perception that different assistant superintendents used different criteria for assigning ratings and a lack of confidence that the rubric did a good job distinguishing effective from ineffective principals or was fair to all principals regardless of the type of school in which they worked.

Principals' opinions about the bonus measure were mixed; most principals did not report negative effects of the bonus on principal morale or on principals' willingness to collaborate, and fewer than half agreed that "[r]ewarding individual principals based on test score gains is problematic because the principal has limited control over student learning in the school." This finding is noteworthy because it suggests that majorities of principals support the idea that principals are responsible for student learning. However, majorities of principals expressed concerns about the fairness and validity of the specific measures used to award bonuses in PPIP. In particular, principals believed that the likelihood of receiving a bonus was related to student characteristics, even though our analyses suggested no such relationship. The bonus also did not appear to serve as a motivator for most principals; more than two-thirds reported that the prospect of earning a bonus did not affect their practices. In addition, most principals were unaware of the premium for working in high-need schools or did not view it as an incentive to work in those schools. Throughout the evaluation, principals consistently said that money did not motivate them to work harder or to change their practices to raise student achievement and that they therefore found the idea of pay for performance problematic.

### **To What Extent Have Principals' Skills and Practices Changed over the Course of the Pittsburgh Principal Incentive Program?**

Principals reported spending increasing amounts of time observing teachers and providing feedback on their instruction as PPIP matured. Curriculum coaches corroborated these reports of increased principal presence in classrooms and reported that teachers in their schools found principals' feedback useful and that principals were effective in the areas of providing professional development opportunities, giving feedback on instruction, and helping teachers use data. Additional relevant evidence was obtained from a teacher survey in the last year of the evaluation, on which large majorities of teachers rated their principals highly as instructional leaders. Moreover, when principal survey respondents were asked to select areas in which their skills had grown the most since PPIP implementation, the most-frequently selected skills were (1) observing in teachers' classrooms and providing feedback and (2) evaluating teachers.

As noted earlier, the fact that PPS was undertaking multiple reforms makes attribution of any changes directly to PPIP impossible, and this problem of attribution was made more challenging by the fact that the district explicitly tried to ensure that its multiple reforms worked together seamlessly. This coherence among reform efforts was likely beneficial for promoting the desired changes, and the data we collected from principals suggest that principals were hearing the message of a unified approach to reform. In particular, throughout the evaluation, principals who participated in interviews were reluctant to attribute changes in their practice to PPIP, especially to the financial incentives. Instead, principals tended to associate changes in their practice with RISE, and they noted that RISE and PPIP reinforced one another in a way that helped them focus on instructional leadership.

### **What Conditions Have Changed at the School and Classroom Levels over the Course of the Pittsburgh Principal Incentive Program?**

Principals reported becoming more-active users of data for decision making and facilitating such data use among their staffs, and our data from coaches and teachers suggest that principals were providing time, structure, and direct support for instructional data use. Staff at the schools where we conducted interviews reported increased frequency and depth of instructional data use over the four years of PPIP implementation. More generally, majorities of principals and coaches reported seeing specific evidence that three major instructional improvement strategies—principal feedback, site-specific professional development, and use of data—led to specific changes in classrooms. In addition, principals and coaches who participated in interviews in study year 4 described improvements in teaching that included improved questioning techniques, more-effective data use, and increased collaboration among teachers around instruction. Principals and coaches also described increased levels of student engagement—specifically, that students were taking ownership of their own learning, often because

the increased data use in the building extended to students and made them more aware of their own progress and where they needed to improve.

### **How Did Principals Perform on the Rubric and Bonus Measures, and How Was Performance Related to Principal Mobility?**

Average principal performance on the rubric remained steady over time, with almost all principals being assigned the highest two out of four categories (proficient or accomplished) on almost all standards on the rubric. On average, principals performed most poorly on the standard related to creating a culture of teaching and learning, which included such activities as data use, curriculum implementation, and differentiated instruction. In interviews, principals said they had made progress on data use, but many said they still struggled with becoming familiar enough with all the curricula their teachers used that they could provide substantive support in that area. Our analysis of the rubric revealed that scores on the individual standards and components were correlated, and the rubric appeared to measure a single construct related to principal leadership. We observed some differences in rubric performance across school type and across schools serving students with different characteristics, although the exact nature of these differences changed from one year to the next.

Performance on the bonus measure was also relatively constant over time and was unrelated to the characteristics of the students in the principal's school. This finding is important because principals expressed concerns that the bonus might favor one type of school (e.g., those serving high-income students) over another, but we saw no evidence of this. This stability in average performance on the bonus measure suggests that an achievement bonus program can be designed so that bonuses change to reflect changes in achievement growth but do not change so much as to create the risk of an overwhelming, unanticipated financial burden. Although the bonus formula was based on absolute rather than relative performance measures, the design process took into consideration prior observed variation in test score gains to limit the risk of very high average bonuses.

We found some evidence that the skills and practices measured by the rubric are associated with improved student achievement. In the final year of the evaluation, mathematics achievement growth showed a statistically significant positive correlation with the first three rubric standards and with the total score. In earlier years, there was evidence of a positive correlation between growth in rubric scores and subsequent achievement growth. Together, these findings provide some evidence that the rubric provided a measure of practices and skills that are associated with principal effectiveness. This relationship is foundational to any program that is trying to improve achievement growth by evaluating and improving practice and should be continually monitored.

Our analysis of mobility showed that the percentage of principals who remained at their schools stayed fairly constant throughout PPIP. Although we did not find sig-

nificant differences in prior achievement bonuses by move type (i.e., whether the principal left the district, switched schools, retired, or moved for some other reason), the patterns were suggestive. For example, principals who moved into administrative positions at the central office level had higher-than-average achievement bonuses prior to their move than other principals, and those who moved from a principal to an assistant principal position had lower bonuses. This finding is consistent with the hypothesis that high-performing principals would be likely to receive promotions whereas low-performing principals would be counseled into positions that will offer them additional experience to improve their practices. Principals who left the district without retiring had slightly lower achievement than the overall average, a pattern that, if it continues, will lead to a gradual increase in the average performance of principals in PPS. The mobility analysis also showed that principals who moved to higher-need schools had earned higher bonuses than other principals before the move but that high-need schools experienced greater principal turnover than other schools. High levels of turnover may or may not be considered problematic; if the departing principals are replaced by more-effective leaders, turnover could be desirable, at least in the early stages of a reform like PPIP.

### **How Did Student Achievement Change During the Course of the Pittsburgh Principal Incentive Program, and How Did Racial and Socioeconomic Gaps Change?**

Student achievement growth in PPS exceeded that of the rest of the state in three out of four years of PPIP implementation. In year 4, student achievement growth in grades 4–8 in both mathematics and reading reached their highest levels since the beginning of the evaluation. These findings suggest that the implementation of PPIP was accompanied by acceleration in achievement growth, consistent with the district's goal of promoting more-effective school leadership as a means of improving student achievement.

Race and poverty achievement gaps increased over the course of PPIP implementation when measured using scale scores for the same set of students over time. However, there is suggestive evidence that achievement growth among the lowest-scoring students and at the most-disadvantaged schools was beginning to increase. In the final year of the evaluation, previously low-scoring students experienced greater achievement growth than in prior years, which is consistent with a design feature of the bonus formula that rewarded gains at the low end of the distribution more than comparable gains at the high end. Furthermore, we found evidence that achievement growth in high-need schools increased following the implementation of the high-need bonus premium, suggesting that the premium may be promoting increased achievement growth at the most-disadvantaged schools. These findings suggest that PPIP can be an important part of the district's efforts to reduce the achievement gap.

## Recommendations

The findings presented in this report do not provide definitive evidence regarding the effects of performance-based evaluation and compensation for principals, so they do not support recommendations regarding whether or not districts or states should adopt such policies. Nonetheless, these findings are potentially useful not only for helping PPS as it continues to implement its ambitious reform agenda but also for informing other districts, states, charter management organizations, and other education agencies that are developing new principal evaluation and compensation systems. We provide several recommendations that PPS and other entities might consider as they develop or revise principal evaluation, support, and compensation policies.

### Recommendations for Evaluation System and Measure Development

*Consider incorporating a range of measures into the evaluation system, including measures that reflect input from a variety of stakeholders.* Teachers, coaches, and other school staff can provide an important perspective on principals' leadership quality and can be useful for helping districts understand whether principals' instructional improvement efforts are leading to the desired effects on schools. Information from other stakeholders, such as parents and students, could also prove helpful for understanding principals' performance. These data could be collected in a variety of ways but, for the purposes of formal evaluation, should be collected consistently across schools and ideally in a way that permits quantification of the information. It is critical that any measure used for the purpose of evaluating principals be carefully piloted and subject to an investigation of the validity of that measure for that specific purpose; instruments that work well for providing formative feedback, for instance, could provide misleading information or be subject to score corruption if used for high-stakes evaluation purposes.

*Gather evidence of validity, reliability, and fairness of the system throughout the implementation of the system, not just at the beginning.* The changes we observed in the correlations between the bonus and rubric measures illustrate how the characteristics of scores can change over time. Districts should continually gather evidence to identify changes in rater agreement, relationships among measures, relationships with external criteria, and fairness to all participants. Alignment of the system with broader district or state goals should be a focus of these ongoing investigations.

*Take steps to ensure consistency in application of rubrics across evaluators.* Principals need assurance that the method used to assign ratings is fair to all principals, regardless of what type of school they lead or who their evaluator happens to be. Although traditional methods for measuring interrater agreement are difficult to apply in some systems (e.g., in PPIP, each rater works with only a certain type of school, so it is difficult to obtain multiple ratings for one principal), other approaches could help promote consistency. These include calibration meetings at which the evaluators rate evidence and discuss their rating criteria, comparisons of the distributions of scores produced

by each evaluator to help them identify the extent to which their ratings differ in magnitude or in variability, and examination of the narrative comments that evaluators produce. Regardless of the approach taken, the district or state adopting the evaluation system should let principals know that it is taking steps to improve consistency of evaluation and should provide professional development to raters to promote high-quality evaluation.

*On measures of principal practice, develop a scale that differentiates performance at all points along the distribution.* Because the vast majority of scores assigned to principals were high, the four-point rating scale used in PPIP offered little opportunity to distinguish between the most-outstanding principals and those who are performing well but not at the highest level. Developing a more expansive rating scale that permits fine-grained distinctions at the high end of the distribution could help districts and states ensure that they are encouraging and rewarding the most-effective principals and could provide information that is useful for identifying principals who could serve in mentoring roles. It could also support targeted professional development for principals. One way to expand the rating scale is to award higher scores to principals who not only perform well themselves but who also help promote the skills and knowledge of others with whom they work.

*Involve all stakeholders in any reviews and redesigns of measures used in evaluation systems.* Incentive pay systems designed to reward complex changes in behavior are, by definition, complicated and often controversial. It is important that districts cultivate shared goals among stakeholders, such as parents, teachers, principals, and central office staff. Once shared goals are established, measures can be reviewed based on their ability to attain these goals.

*Monitor racial and socioeconomic achievement gaps using student-level growth throughout the achievement distribution.* Given the importance that many districts have placed on the goal of reducing achievement gaps, it would be worthwhile to monitor progress toward this goal using multiple metrics. The practice of tracking changes in percentage proficient across cohorts provides useful information, but it may lead to distorted impressions regarding the nature and extent of improvement among individual students who are performing at different points in the achievement distribution. By also examining student-level growth in scale scores for students who remain in a district for at least two consecutive years, the district could provide richer and more-accurate information to inform its own decision making and to enhance public understanding of the district's progress.

## **Recommendations for Implementation**

*Align the elements of a performance-based compensation system, including support and criteria for evaluation, with the district's approach to improving teaching and learning.* Principals who are participating in multiple, simultaneous reforms are more likely to support the reforms and respond effectively when the reforms are well connected and

aligned with a common set of goals. Districts undertaking reforms like PPIP should consider the extent to which evaluation criteria, professional development, and other elements of the reform support or conflict with other key initiatives.

*Devise a communication strategy that provides clear, timely, and ongoing information to help principals understand the evaluation measures and the steps the district took to ensure their validity.* Principals' concerns and lack of understanding of some aspects of PPIP suggest that any effort by a district or other entity to adopt performance-based compensation should be accompanied not only by efforts to establish the validity and reliability of the measures but also by a communication strategy that provides clear, timely, and ongoing information to stakeholders. A comprehensive communication strategy should involve multiple vehicles of communication, including large-group meetings, as well as individualized interactions and stakeholder involvement in decisions about changes to the program.

*Provide principals with concrete tools for accomplishing the instructional leadership tasks (especially observing and providing feedback on instruction) encouraged by the compensation system.* In PPIP, such resources as professional development, targeted support from supervisors, and a clear set of standards communicated by the rubric, were perceived by principals as contributing to their professional growth. The positive perceptions about the utility of the DPG projects suggest that this type of activity could be a valuable tool, particularly for helping more-experienced and more-accomplished principals enhance their knowledge and skills in a particular area.

*Help principals find the time needed to engage in the practices promoted by the initiative.* If the initiative encourages principals to spend more time on specific tasks, such as supporting teachers, it is important to help them find that time. Districts could encourage discussion of tasks on which principals might spend less time and the most-effective ways to incorporate delegation in their leadership approach. Districts could also create structures that facilitate delegation. The ability to cultivate leadership among school staff is an important element of effective leadership and could be directly incorporated into the evaluation and support systems, particularly for experienced principals.

*Assess the extent to which principal mobility leads to improved access to effective principals at high-need schools and to higher levels of principal effectiveness overall.* We found that high-need schools experienced greater principal mobility than other schools, but also that principals who switched schools and moved to higher-need schools had higher prior performance on the bonus measure than those who moved to lower-need schools. It is not clear whether the higher mobility levels in high-need schools will lead to disruption or to improved educational quality, so districts and states should monitor mobility over time to determine whether equity and overall effectiveness are improving.

## Conclusion

The implementation of PPIP during the period of the TIF grant was accompanied by changes in the practices, skills, and attitudes of principals and other school and district staff. Each year, the district modified elements of the program in response to challenges it encountered. The district's experiences can be informative for efforts to establish new evaluation and compensation systems for principals in other districts or in states, charter management organizations, or other education agencies. Some of the recommendations presented in this summary stem from the unique context in which PPIP was operating, but most of them are likely to be applicable to some degree in other contexts and can serve as a resource for future principal evaluation reforms.