Changing the Principal Supervisor Role to Better Support Principals

Evidence from the Principal Supervisor Initiative

Ellen B. Goldring, Melissa A. Clark, Mollie Rubin, Laura K. Rogers, Jason A. Grissom, Brian Gill, Tim Kautz, Moira McCullough, Michael Neel, and Alyson Burnett

VANDERBILT PEABODY COLLEGE

Mathematica

JULY 2020

Changing the Principal Supervisor Role to Better Support Principals

Evidence from the Principal Supervisor Initiative

Ellen B. Goldring, Melissa A. Clark, Mollie Rubin,
Laura K. Rogers, Jason A. Grissom, Brian Gill, Tim Kautz,
Moira McCullough, Michael Neel, and Alyson Burnett

VANDERBILT PEABODY COLLEGE

Mathematica

Commissioned by:

The Wallace Foundation®

JULY 2020

ACKNOWLEDGMENTS

This report would not have been possible without the contributions of many individuals. First and foremost, we are grateful for the cooperation of the central office staff, principal supervisors, and principals from the six school districts participating in the Principal Supervisor Initiative. We especially want to thank the district staff who facilitated our work: Kasey Mengel in Baltimore City Public Schools, Veda Hudge in Broward County Public Schools, Andrea Foxx in Cleveland Metropolitan School District, Ruth Wright in Des Moines Public Schools, Kimberly Dueñas in Long Beach Unified School District, and Jennifer Fair in Minneapolis Public Schools.

We also thank The Wallace Foundation staff who have provided essential information about the initiative and thoughtful insights throughout the study. We particularly appreciate support and insights into the initiative from the Education Leadership Team, including Jody Spiro, Aiesha Eleusizov, Rochelle Herring, and Nicholas Pelzer, and expert guidance on the study from Elizabeth Ty Wilde, Ed Pauly, Rachel Hare Bork, Andrew Cole, and Jessica Schwartz.

We are grateful for input on the study design provided by members of the study's technical working group, which included Richard Murnane, Jeffrey Smith, Petra Todd, and Brenda Turnbull. The study has benefited greatly from their expertise. Andy Porter provided valuable guidance on the analysis of data from the Vanderbilt Assessment of Leadership in Education (VAL-ED). We also appreciate helpful input on the analysis from Susan Gates, Matthew Baird, and Benjamin Master.

The study has also benefited from the contributions of many people at Vanderbilt University and Mathematica Policy Research. At Vanderbilt, Tammy Eidson provided crucial logistical support. From Mathematica, Charles Tilley and Kerry Schellenberger provided expert programming assistance, Ruth Neild provided insightful comments on the report, Brigitte Tran helped develop report graphics, and Sheena Flowers provided expert production support.

CONTENTS

TABLES

FIGURES

EXECUTIVE SUMMARY

In 2014, The Wallace Foundation launched the Principal Supervisor Initiative (PSI), a four-year, $24 million effort to redefine principal supervision in six urban school districts. The PSI aimed to help districts overhaul a position traditionally focused on administration, operations, and compliance to one dedicated to developing and supporting principals to be effective instructional leaders in their schools.

The PSI was motivated by The Wallace Foundation's longstanding commitment to improving students' academic achievement by strengthening the quality of educational leadership. Research suggests that strong principals are a significant factor in teachers' instructional quality and school success. Principal supervisors are a potential point of leverage for supporting and developing principals' effectiveness, but little is known about the effectiveness of this approach. The overarching hypothesis of the PSI was that changing the role of principal supervisors from overseeing operations to developing principals' instructional leadership practices could drive improvement in principal effectiveness.

Six districts participated in the PSI:

1. Baltimore City Public Schools, Maryland[a]
2. Broward County Public Schools, Florida
3. Cleveland Metropolitan School District, Ohio
4. Des Moines Public Schools, Iowa
5. Long Beach Unified School District, California
6. Minneapolis Public Schools, Minnesota

[a] DeKalb County School District (Georgia) was an initial participant in the PSI but withdrew after the first year, following a change in district leadership. Baltimore City Public Schools joined the initiative after DeKalb withdrew.

The PSI includes five core components:

1. Revise the principal supervisors' job description to focus on instructional leadership
2. Reduce principal supervisors' span of control (the number of principals they oversee) and change how supervisors are assigned to principals
3. Train supervisors and develop their capacity to support principals
4. Develop systems to identify and train new supervisors (succession planning)
5. Strengthen central office structures to support and sustain changes in the principal supervisor's role

The Wallace Foundation commissioned an independent study of the PSI. The study, conducted by researchers from Mathematica Policy Research and Vanderbilt University, documents districts' experiences implementing the PSI and its effect on teachers' perceptions of principals' performance. An initial report, "A New Role Emerges for Principal Supervisors: Evidence from the Six Districts in the Principal Supervisor Initiative" (Goldring et al. 2018), traced the implementation of the PSI from its inception in August 2014 through the first three years of the initiative. This second report describes central offices', principal supervisors', and

principals' experiences with the PSI, as well as the PSI's effects on teachers' perceptions of principals' performance over the four years of the initiative. A third report will compare principal supervision in the PSI districts with that in other urban districts throughout the country.

A. Objectives

This report describes districts' experiences with the PSI across the years of the initiative and analyzes its effects on teachers' perceptions of principals' performance. Specifically, the report addresses the following questions:

- What is the theory of change for revising the principal supervisor's role?
- What were districts' and principal supervisors' experiences with the PSI?
- What were principals' experiences with the PSI?
- How did the initiative affect teachers' perceptions of principals' performance, and how did teachers' perceptions of principals' performance and the effects of the PSI vary across districts and principal supervisors?
- How do key components of the PSI relate to changes in teachers' perceptions of principals' performance?
- How did the principal supervisor role in PSI districts compare with that in other urban districts?

The report also discusses the sustainability of changes made during the PSI and lessons learned for revising the principal supervisor role.

B. Methods

The report is based on analyses of data from multiple sources:

- **Site visits and interviews.** We conducted three rounds of semistructured interviews with the superintendent and a sample of central office personnel, principal supervisors, and principals in each of the six PSI districts over the course of the initiative. Through the interviews, we aimed to understand districts', principals', and supervisors' ongoing experiences as the district changed the principal supervisor role.

- **Surveys.** We administered surveys to all principal supervisors and principals in the PSI districts three times during the initiative. The survey data provided the perspectives of a broader range of supervisors and principals regarding each of the PSI components. We also administered a survey to principal supervisors in a sample of other urban districts that were not part of the PSI during the final year of the initiative.

- **Teacher ratings of principals' performance.** We measured teachers' perceptions of principals' performance using teachers' ratings of principals from the Vanderbilt Assessment of Leadership in Education (VAL-ED), a principal evaluation tool focused on the same types of instructional leadership practices and competencies emphasized by the PSI. We used VAL-ED data from the 2013–2014 through 2017–2018 school years to analyze the PSI's effects on teachers' perceptions of principals' performance.

Our analyses of data from site visits, interviews, and surveys focused on districts', supervisors', and principals' experiences, successes, and challenges throughout the initiative; lessons learned; and plans for sustaining the accomplishments of the PSI after the initiative ended. We integrated the descriptive survey results with the qualitative data.

To provide rigorous evidence of the effects of the PSI on teachers' perceptions of principals' performance, we first used propensity score matching to identify a comparison group of schools that was similar to the schools in the PSI districts before the start of the initiative. We then tracked and compared principals' VAL-ED scores in the two groups over time to estimate the PSI's effects on teachers' perceptions of principals' performance. In addition, we examined how teachers' perceptions of principals' performance and the PSI's effects varied across the PSI districts and the schools in these districts. We also used survey data to construct measures of PSI implementation factors, including implementation of specific components of the PSI; principal supervisors' time spent on instructional leadership; and supervisors' effectiveness. We examined how these factors related to the PSI's effects on teachers' perceptions of principal's performance.

Research question	Data and methods used
What is the theory of change for revising the principal supervisor's role?	• Synthesis of districts' approaches to implementing PSI components based on data from site visits and interviews
What were districts' and principal supervisors' experiences with the PSI?	• Descriptive analyses of data from site visits, interviews, and surveys
What were principals' experiences with the PSI?	• Descriptive analyses of data from site visits, interviews, and surveys
How did the PSI affect teachers' perceptions of principals' performance? How did teachers' perceptions of principals' performance and the effects of the PSI vary across districts and principal supervisors?	• Comparison of teachers' perceptions of principals' performance (as measured by teachers' ratings from the VAL-ED) for schools in PSI districts and similar schools in non-PSI districts • Analysis of variation in teachers' perceptions of principals' performance and PSI effects
How do key components of the PSI relate to changes in teachers' perceptions of principals' performance?	• Examination of correlations between school-level effects and implementation factors
How did the principal supervisor role in PSI districts compare with that in other urban districts?	• Decriptive analysis of data from principal supervisor surveys in PSI districts and a sample of other urban districts

C. District context in the final year of the PSI

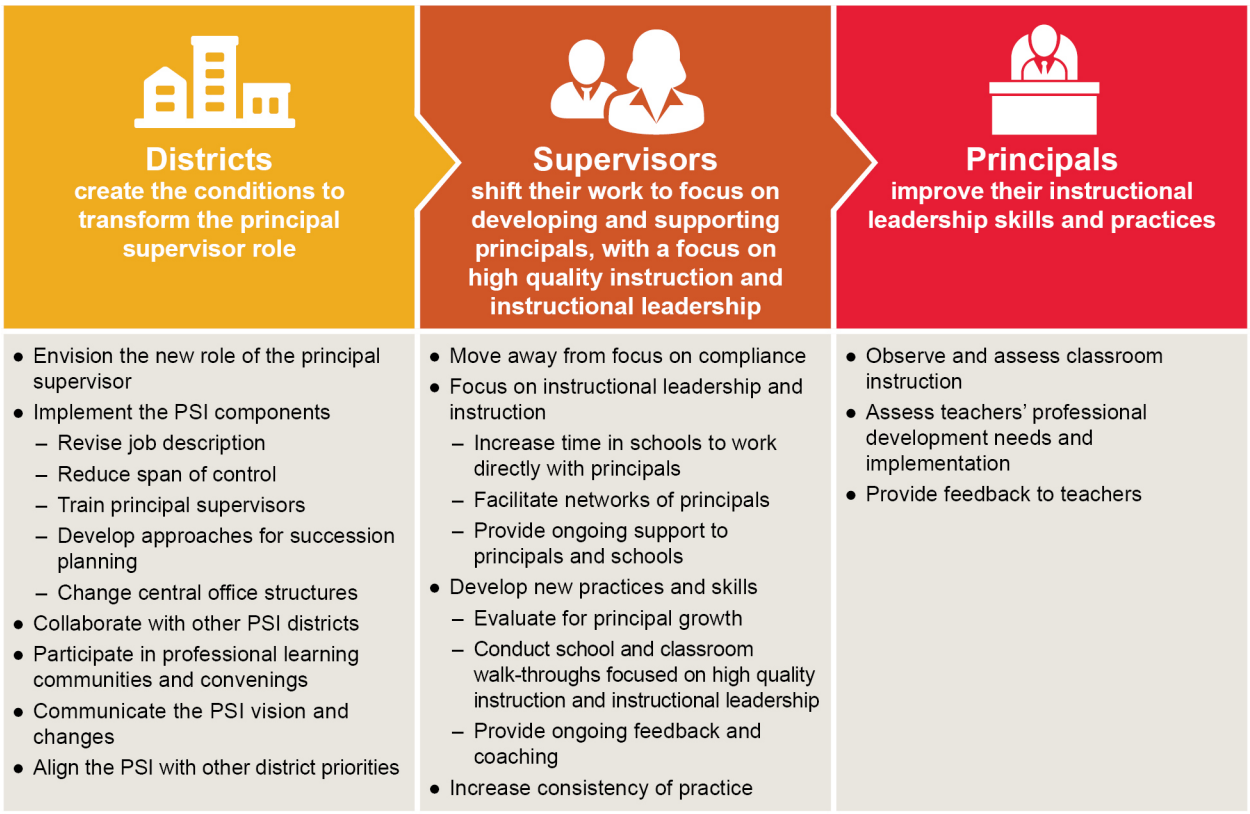
Each district's particular circumstances in the final year of the PSI influenced its experiences and progress with the initiative. One district faced budget shortfalls that required cuts to many central office departments and positions, along with the elimination of some principal supervisor positions. Some districts experienced leadership changes or central office reorganization, along with changing district priorities. Other districts, in contrast, benefited from longstanding stability in central office leadership and strong school board support. All districts faced challenges as they continued to focus on how best to meet the needs of low-performing schools in complex urban environments.

D. Findings

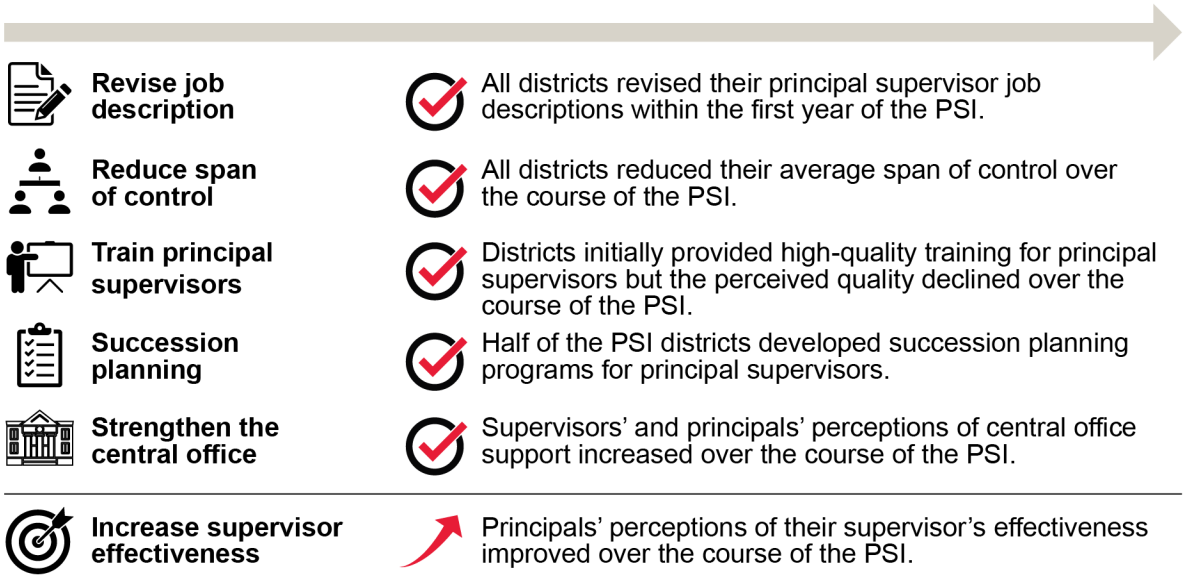
The PSI brought substantial changes to all six districts. The initiative was successful in changing the principal supervisor role to one that focused on providing support and guidance to schools and principals. Principals' perceptions of their supervisors' effectiveness improved over the course of the initiative. Districts restructured their central offices to better support schools and meet their needs. Changes districts made during the PSI led to a shared understanding of the principal supervisor role and equipped principal supervisors to better support principals. However, the PSI did not improve teachers' perceptions of principals' performance, on average.

1 **The PSI focused on district-level changes to redefine the principal supervisor role to support and improve principals' performance.** Although the PSI laid out a broad approach to overhauling the principal supervisor role through its five core components, it did not directly define the specific approaches districts should take to achieve this goal. Rather, it encouraged participating districts to implement the components in ways that aligned with their particular contexts. The theory of change emerged over the course of the initiative as districts adapted the PSI components to define, change, and support the supervisor role according to their unique contexts (Figure 1).

Figure 1. The theory of change that emerged as districts revised the principal supervisor role



2 Districts and supervisors made substantial changes as a result of the PSI. The PSI led to district-level changes in central office culture, structures and support for schools.



- **All districts revised their principal supervisor job description early in the initiative.** The job description served as the basis for communicating the change in the principal supervisor role to staff and community members and helped drive the district's vision for principal supervision. The job description served as a guide for districts over the course of the PSI.
- **PSI districts reduced their average span of control from 17 to 13 principals.** Supervisors with smaller spans of control spent more time meeting with their principals. They also increased their emphasis on instructional leadership practices during school visits, including practices to support data use, teacher feedback, and classroom visits. The PSI districts also reduced principal supervisors' outside responsibilities and other formal central office roles, which allowed the supervisors to place a more consistent focus on supporting principals. All supervisors spent the majority of their time working with principals one-on-one or in group meetings.
- **In the early years of the initiative, the PSI districts worked hard to provide dedicated training for principal supervisors.** Districts sought to clearly define and strengthen supervisors' understanding of high-quality instruction and instructional leadership. They worked with external technical assistance providers to deliver common tools and training to principal supervisors and other central office staff. Districts also aimed to provide a standard of support to principals and schools that was aligned with the district's vision of high-quality instruction, instructional leadership, and school support. In the later years of the PSI, dedicated training declined in both quality and quantity in some districts. In these districts, dedicated time for professional learning and continuous development was often overtaken by administrative items and information sharing.

- **Half of the PSI districts developed specific programs for succession planning for principal supervisors.** Districts also became more systematic and comprehensive about their approaches to screening and hiring new principal supervisors, increasing the rigor of their selection processes. However, districts with aspiring principal supervisor programs faced a challenge in trying to balance the small number of supervisor vacancies each year with a group of aspiring supervisors who had completed a program but yet faced few to no openings in the district. Over the course of the initiative more supervisors became aware of open principal supervisor positions because of aspiring supervisor programs and opportunities. Some districts also adapted their aspiring supervisor programs to use for filling other central office leadership positions.
- **Districts modified central office structures, roles, and culture to better support principal supervisors.** Districts developed new communication systems and approaches for coordinating with supervisors and schools. They continued to bolster structures they created early on, such as cross-departmental liaisons and central office support teams, to orient the central office toward supporting principals in schools. Supervisors increasingly collaborated with other district departments to plan and coordinate principal professional development and support. Supervisors' and principals' perceptions of central office support increased steadily over the course of the PSI (Figure 2).
- **Principal supervisors changed their practices, and practices became more consistent within some districts.** Supervisors began to develop a shared professional identity beyond middle management and a universal set of norms and skills to guide their practice. They implemented specific coaching models, using protocols for school walk-throughs, and providing feedback to principals. Supervisors worked with principals to help them develop effective teachers through classroom observations and teacher feedback. Some districts strived to improve consistency by sharing common goals for supervisors, standardizing the focus of school visits, adopting common tools to guide supervisors' interactions with principals, working to calibrate evaluation ratings, and setting expectations for the amount of support supervisors should provide to principals and schools.
- **Despite widespread progress, districts and supervisors faced challenges as they implemented the new supervisor role**. The quality of PSI implementation varied both across the six districts and within each district. Many central office level changes, such as increasing cross-departmental communication and responsiveness to school needs, challenged long-standing organizational culture and context within the district. Additionally, supervisors sometimes found the new role itself to be demanding compared with the previous, compliance-oriented role. In some districts, budget shortfalls and changes in district leadership made it difficult to sustain changes to the principal supervisor role, and the role started to revert back to a focus on compliance.

Figure 2. Principals' perceptions of central office support improved

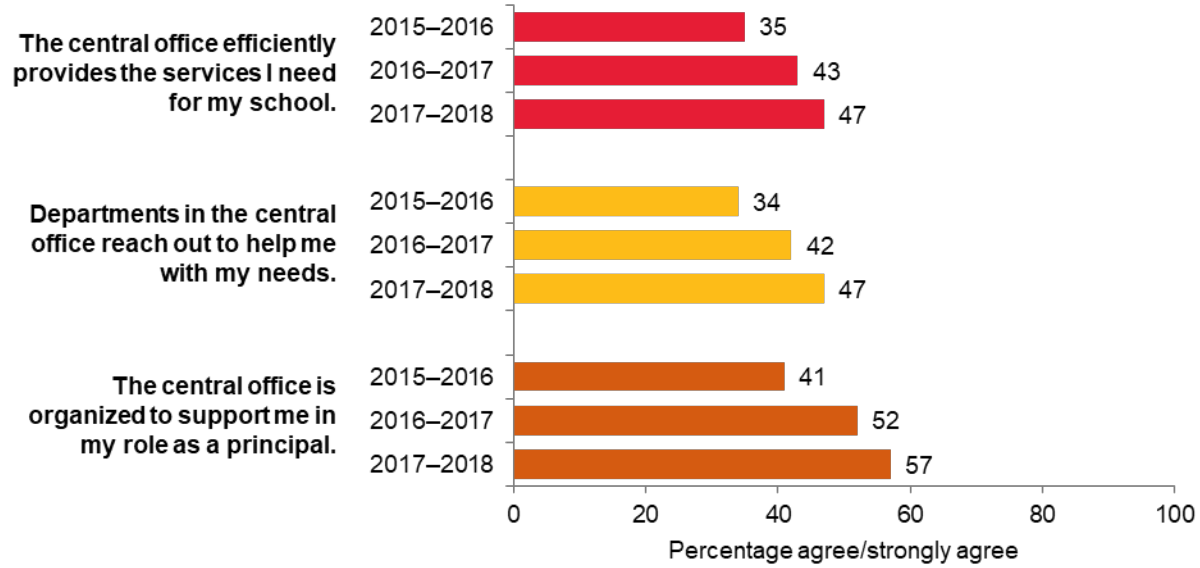


Figure reads: Thirty-five percent of principals agreed or strongly agreed that departments in the central office efficiently provide the services they need for their schools in 2015–2016.

Source: Principal surveys, 2016 (*N* = 635), 2017 (*N* = 639), and 2018 (*N* = 606).

Note: The survey prompt read, "Please indicate to what extent you agree or disagree with the following statements." Changes across the three years are statistically significant at the 5 percent level for all three items.

3 **Principals' perceptions of their work with their supervisor and their supervisor's effectiveness improved.** Principals noticed many changes in their work with their supervisor and their own leadership capacity over the course of the PSI.

- **Principals' expectations about their work with their supervisors changed.** Principals' understanding of effective support changed during the PSI, as did the types of relationships they wanted to have with their supervisors. Principals came to expect a consistent relationship with their supervisors that included coaching, feedback, deeper professional development, and stronger supports for and expertise about instruction.
- **Principals' perceptions of their supervisor's effectiveness improved over the course of the initiative**. They also reported increased frequency of supervisor implementation of practices encouraged by the PSI, such as coaching, feedback, and data use (Figure 3). Principals recounted, over the course of the initiative, how their relationships with their supervisors had improved because supervisors better understood their contexts and specific needs; deeper relationships of trust and respect developed.
- **Principals' perceptions of their districts' approaches to their evaluation also improved.** Principals noted that deeper relationships with their supervisors helped improve the quality and usefulness of the evaluation feedback they received from their supervisors.

- **Principals reported that they increased their leadership capacity.** Increased support from supervisors facilitated a shift in principals' perceptions of their own capacity to address instructional matters in their buildings. Many principals also noted a shift in their own understanding of their roles as instructional leaders. Specifically, some principals felt more independent, valued, and focused on instruction. Some principals noted that a strong focus on understanding and defining high-quality instruction enabled them to better assess instructional quality and rigor.

Figure 3. Principals' perceptions of supervisors' practices and effectiveness increased

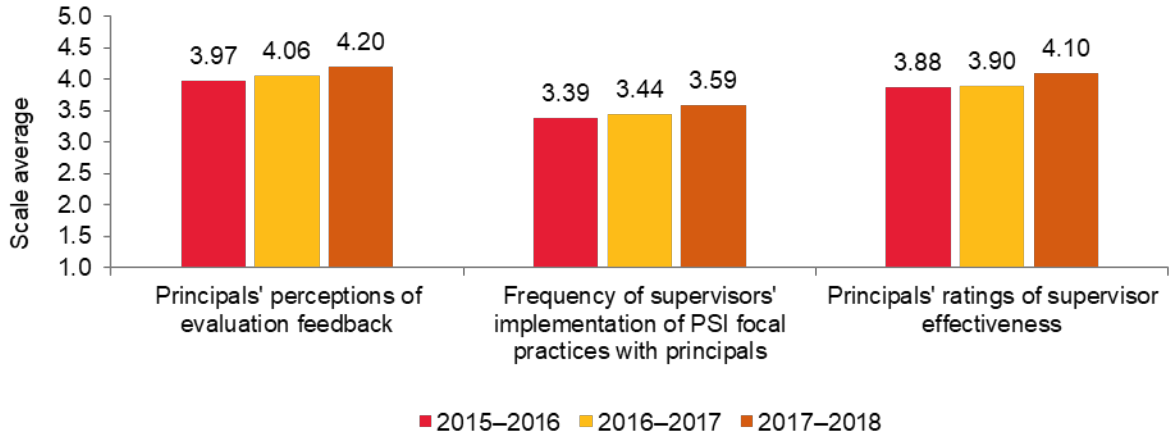


Figure reads: Principals rated the quality of their supervisors' evaluation feedback an average of 3.97 out of 5 in 2015–2016.

Source: Principal surveys, 2016 (*N* = 635), 2017 (*N* = 639), and 2018 (*N* = 606).

Note: All scales range from 1 to 5. See Appendix A for information on scale creation. Changes across the three years were statistically significant at the 5 percent level for all measures.

4 **The PSI did not improve teachers' perceptions of principals' performance.** Teachers' perceptions of principals' performance in the PSI districts remained similar to that in a set of similar comparison schools in districts that did not implement the PSI. The PSI's effects on teachers' perceptions of principals' performance varied across schools, but the effects did not vary systematically across districts or supervisors. This suggests that districts' differing approaches to implementing the PSI and the behaviors of individual supervisors were not driving the differences in effects across principals. It is possible that the PSI affected other aspects of principals' performances not captured by the study, such as principals' ability to hire or retain more effective teachers, or other important outcomes such as principal retention or satisfaction. In addition, the timeframe for the study may have been too short to detect changes in teachers' perceptions of principals' performance resulting from the PSI.

5 **Principals' perceptions of their supervisors' effectiveness and use of certain practices promoted by the PSI had a small positive relationship with the PSI's effects on teachers' perceptions of principals' performance.** This suggests that the supervisor practices emphasized by the PSI—such as helping principals analyze data to make school decisions, focusing the principal's time on instruction and teaching, or providing actionable feedback—could help improve teachers' perceptions of principals' performance. We found

little to no association between the PSI's effects and other implementation factors, including supervisor time spent on instructional leadership, quality of central office support, and supervisor training quality.

6 **The principal supervisor role in the PSI districts differed from that in other urban districts in several key ways, but there were also some important similarities in supervisors' work with principals.** Compared with supervisors in other urban districts, supervisors in the PSI districts had lower spans of control, received more training and mentoring, spent less time on operational issues, and had more favorable views of their districts' principal and supervisor evaluation systems. These differences suggest that the PSI led participating districts to make greater changes to these aspects of principal supervision than the general changes occurring in other urban districts nationwide during the same time frame. However, principal supervisors in PSI and other urban districts spent similar amounts of time working with principals and used similar instructional leadership practices in this work. Principal supervisor standards and other national and local efforts to shift the focus of the principal supervisor role in ways that mirrored some of the changes promoted by the PSI might have driven these similarities.

E. Looking ahead: Sustaining the changes in the PSI districts

For PSI districts that continue to focus on supporting and developing the principal supervisor role, several district actions will be important for sustaining changes implemented during the PSI.

- **Embedding the principal supervisor role within the broader structures and work of the central office.** Some districts connected central office staff and principal supervisors through structures such as cross-departmental and cross-functional teams, training, and meetings. Principal supervisors in many districts took on leadership of districtwide school improvement efforts. These new approaches to central office and supervisor interactions facilitated ongoing communication and relationship building and can continue to build sustainability of the PSI.
- **Communicating the importance of high-quality principal support to stakeholders.** Stakeholders who understood how the PSI aligned to district goals were more supportive of efforts to reallocate roles and resources around principal supervision. Ongoing communication and explanations of the work for department chiefs and school board members can help sustain the achievements of the PSI.
- **Obtaining financial resources to support PSI changes.** Districts that allocated financial resources to fully fund the added supervisor positions and transition from external to in-house training for principal supervisors were most successful in implementing the PSI. Those districts will likely be able to sustain the initiative's momentum.
- **Developing an understanding among senior district leaders that the principal supervisor role is not static.** Several districts planned to revisit and refresh the principal supervisor job description and responsibilities to match evolving district goals for school support, thus ensuring the long-term relevance of the role. This understanding is central to the ongoing goals of the PSI.

- **Championing of the PSI and its vision by the superintendent after the initiative ends.** Districts with longer-standing superintendents who remained engaged with the PSI were well-positioned for sustainability. As the initiative drew to a close, these superintendents could clearly articulate how they planned to continue and scale up the work of the PSI within their districts. Superintendents spoke of ensuring sustainability of the PSI by (1) safeguarding reduced span of control by securing necessary funding and preventing the addition of multiple additional responsibilities, (2) aligning principal supervision with district-wide priorities and initiatives, (3) continuing to signal support of the PSI work to stakeholders, and (4) articulating a commitment to ongoing central office reform.

F. Lessons learned for revising the principal supervisor role

The experiences of the PSI provide lessons for other districts to consider when implementing district-level efforts to revise the principal supervisor role.

- **Obtain buy-in and build awareness across all stakeholders.** District leaders repeatedly referenced the importance of building buy-in and awareness across all stakeholders in the district, including board members, central office staff and principals, to ensure the success of the PSI.
- **Balance supervisors' focus on instructional leadership with the flexibility to meet principal needs.** While a main thrust of the PSI was to focus supervisors on supporting principals' instructional leadership, this was not meant to be supervisors' only focus. Supervisors, principals, and central office staff described the importance of building flexibility into the role to meet the wide-ranging needs of principals and their schools. For instance, principal supervisors described the need to spend far more time on logistical and operational issues as well as instructional leadership with new principals.
- **Invest in selecting and training the best candidates for the principal supervisor position.** While many district officials and supervisors believed prior principal experience to be an asset, leaders in all districts noted that the supervisor role was dramatically different from the principalship and required a diverse set of skills. In addition to identifying high-performing principals as potential supervisors, districts also looked for teaching-oriented individuals who were reflective, good listeners, skilled at data use, and oriented to the district's mission.
- **Consider the trade-offs in different strategies for assigning principals to supervisors.** As districts reduced principal supervisors' span of control, they employed different strategies to assign supervisors to schools. Supervisors and principals valued the convenience of geographic clustering and the opportunity to work with other principals with similar school specialties or demographically similar student bodies. At the same time, some principals in networks of similar schools expressed the desire to have more opportunities to learn from principals who could offer different perspectives.
- **Consider the stability of district context and leadership.** District stability was important for positive PSI experiences. Supervisors and district leaders found it more difficult to address challenging aspects of the PSI when faced with such contextual factors as superintendent changes, turnover of top-level central office leaders, and deep resource and financial constraints.

- **Change structures and transform values, beliefs, and behaviors.** Changing the supervisor role required changes to the structures in the central office. However, it also required adaptive change—or changing values, beliefs, and behaviors in the district. Under the revised supervisor role, while schools had to be accountable to the central office, the central office had to be accountable to schools as well.

G. Questions for further research

Future research should continue to explore the effectiveness of efforts to leverage the principal supervisor role to support and develop principals as instructional leaders. Many of the changes to the supervisor role promoted by the PSI, including reductions in span of control and increases in support and training for principal supervisors, have been adopted to varying degrees by urban school districts throughout country (Cochran et al. 2020). These broader national trends, along with varying approaches to principal supervision across a diverse set of districts, will provide additional opportunities to explore how best to leverage the principal supervisor role to improve principal's performance. For example, research might examine district-level efforts to revise the principal supervisor role in conjunction with implementation of a principal pipeline to develop, select, and support principals. Future research could also examine the effectiveness of such efforts on a broader set of outcomes beyond teachers' perceptions of principals' performance, including principal retention, teacher retention and satisfaction, and student achievement. Finally, future research might continue to explore the relationship between teachers', principals', and supervisors' perceptions of principals' performance.

I. OVERVIEW OF THE PRINCIPAL SUPERVISOR INITIATIVE

In 2014, The Wallace Foundation launched the Principal Supervisor Initiative (PSI), a four-year, $24 million effort to redefine principal supervision in six urban school districts. The PSI aimed to help districts overhaul a position traditionally focused on administration, operations, and compliance to one dedicated to developing and supporting principals to improve instruction in schools. This overhaul involved direct changes to the principal supervisor role as well as changes to central office structures and policies to better support supervisors in their work with principals. The overarching hypothesis of the PSI was that changing the role of principal supervisors from overseeing operations to providing instructional leadership could drive improvement in principal effectiveness. Improved principal effectiveness could, in turn, be an important lever for improving instruction and, ultimately, student performance.

The Wallace Foundation commissioned an independent study of the PSI to share lessons from the initiative with school districts, education practitioners, policymakers, and researchers. The study, conducted by researchers from Mathematica and Vanderbilt University, has two interrelated aims: to document districts' experiences implementing the PSI and to determine its effect on teachers' perceptions of principals' performance. More broadly, the study provides an opportunity to learn whether a district-level intervention that does not directly interact with principals can improve principals' performance.

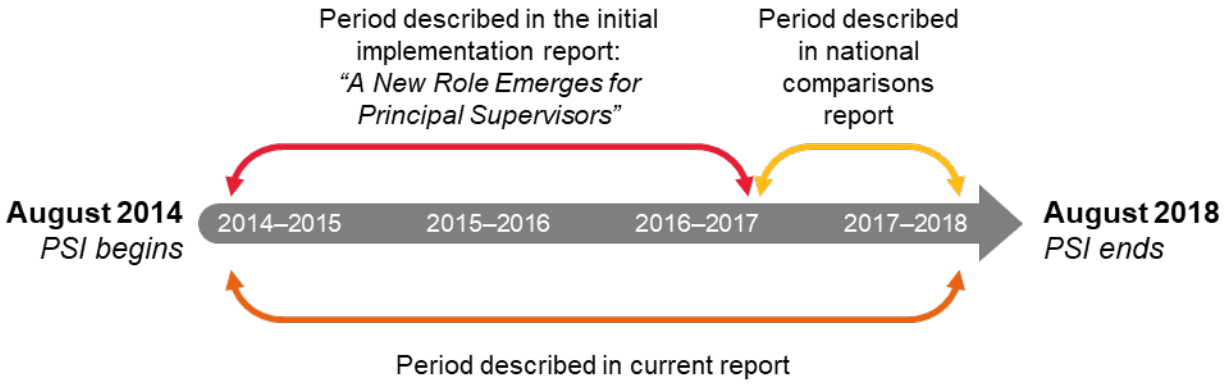
Six urban districts participated in the PSI:

1. Baltimore City Public Schools, Maryland[a]
2. Broward County Public Schools, Florida
3. Cleveland Metropolitan School District, Ohio
4. Des Moines Public Schools, Iowa
5. Long Beach Unified School District, California
6. Minneapolis Public Schools, Minnesota

[a] DeKalb County School District (Georgia) was an initial participant in the PSI but withdrew after the first year, following a change in district leadership. Baltimore City Public Schools joined the initiative after DeKalb withdrew.

An initial report, "A New Role Emerges for Principal Supervisors: Evidence from the Six Districts in the Principal Supervisor Initiative" (Goldring et al. 2018), traced the implementation of the PSI from its inception in August 2014 through spring 2017. That report described the districts' individual approaches to implementing the PSI over the first three years of the initiative, as well as their accomplishments and challenges. This second report, "Changing the Principal Supervisor Role to Better Support Principals: Evidence from the Principal Supervisor Initiative," describes central offices', principal supervisors', and principals' experiences with the PSI, as well as the PSI's effects on teachers' perceptions of principals' performance over the four years of the initiative. A third report will compare principal supervision in the PSI districts with that in other urban districts throughout the country in the 2017–2018 school year (Figure I.1).

Figure I.1. Time periods covered in the three study reports



A. The PSI: Background

The PSI evolved from The Wallace Foundation's longstanding commitment to improving students' academic achievement by strengthening the quality of educational leadership. Research suggests that principal effectiveness is a significant factor in school success. For example, schools with effective leaders have more satisfied teachers, lower rates of teacher turnover, more positive learning climates, greater parent engagement, and, ultimately, higher student achievement (Boyd et al. 2011; Grissom et al. 2015; Grissom and Loeb 2011; Hallinger et al. 1996; Sebastian and Allensworth 2012). Research also shows that the principal role has shifted over the past several decades. Increasingly, school districts expect principals to improve their teachers' performance through observations, feedback, and other forms of instructional leadership (Neumerski et al. 2018).

Although there is a large body of research on effective school leadership, less is known about how to develop and support leaders to become more effective (Goff et al. 2014). Many efforts to improve principal effectiveness involve direct work with principals, through individualized coaching, professional development, or peer mentoring. However, rigorous research on these approaches is limited, and the few rigorous studies show that they are not consistently effective in improving principals' leadership practices or student achievement (see, for example, Jacob et al. 2014).

More recently, however, districts themselves are becoming a lever to improve principal effectiveness through changes to district policies, structures, and practices. These efforts follow research that suggests that school district policies, structures, and practices are related to district effectiveness (Anderson and Young 2018a). This research defines district effectiveness as "the ability of a district to achieve the mission of delivering high quality and equitable educational experiences for each student" (Anderson and Young 2018a, p. 2). The factors that are related to district effectiveness include such practices as aligning structures, resources, and personnel policies; having a districtwide focus on student achievement; and building and maintaining good internal and external communications and relations. However, much of this research base is descriptive and correlational (Anderson and Young 2018b). The research does not provide much insight into what effective districts do to develop and support effective leadership.

Strategic efforts by districts to improve school leadership effectiveness are varied. One such districtwide effort is the implementation of principal pipeline programs to identify and prepare aspiring principals, as well hiring principals and giving them on-the-job support and evaluation. Pipeline programs typically embed training for the principalship in the aspiring principal's job (typically as an assistant principal or teacher leader) and provide comprehensive field experiences or internships with ongoing feedback. Recent evaluations of the New Leaders Aspiring Principals Program (Gates et al. 2019a) and The Wallace Foundation's Principal Pipeline Initiative (Gates et al. 2019b) found that these pipeline programs had positive effects on student achievement and principal retention. These recent studies suggest that district interventions can improve the quality and effectiveness of school leaders.

The PSI was a deliberate effort by The Wallace Foundation to further examine the role that districts can play in supporting and developing effective principals. Principal supervisors, sometimes called area superintendents, assistant superintendents, network leads, directors of school support, or other titles, are the central office staff who directly oversee and evaluate principals. The principal supervisor is a natural locus for supporting and developing principals' instructional leadership skills, and research suggests that improving the effectiveness of principal supervisors could be an important lever for improving principals' effectiveness (Corcoran et al. 2013). Furthermore, studies of district reform strategies suggest that principal supervisors can play a key role in helping principals develop instructional leadership skills and that district investment in training principal supervisors can help propel school improvement efforts (Marsh et al. 2005). The PSI sought to improve principals' effectiveness both through direct training of principal supervisors and through changes to central office structures and policies to better support principal supervisors in their work with principals. The initiative provided $18 million directly to the six districts (an average of $3 million per district), along with an additional $6 million for technical assistance, professional learning community meetings, and research and evaluation.

B. The core components of the PSI

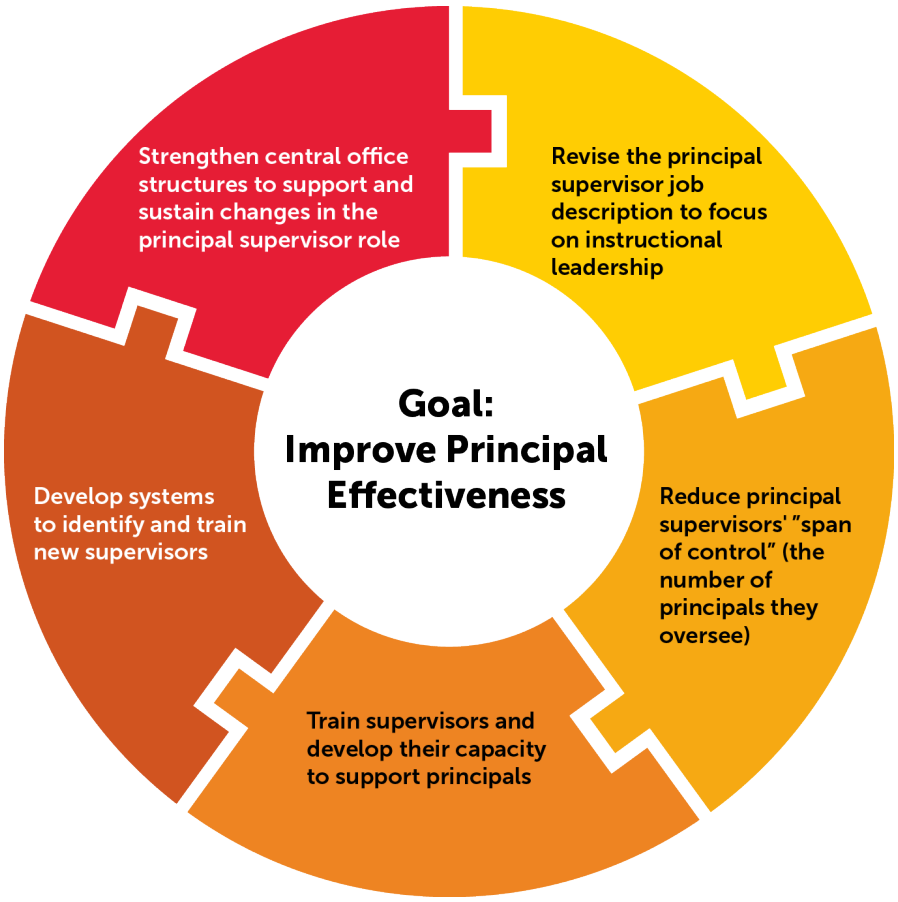
The PSI consists of five core components (Figure I.2) that districts implemented according to their local contexts and needs.

Revising the principal supervisors' job description to focus on instructional leadership. Principals need support from supervisors to act primarily as instructional leaders, rather than as building managers. Traditionally, however, central offices required principal supervisors to focus on compliance, such as ensuring that principals had submitted appropriate forms for budgeting and state accountability, checking on the completion of school improvement plans, and monitoring whether Individualized Education Plans were up to date. This left principal supervisors with little time to provide support for instructional leadership. Thus, one key component of the PSI is to help districts reorient expectations for supervisors' work to focus on instructional leadership in schools. Revising the job description is the first step for district leaders to redefine and codify their vision for the principal supervisors' role.

Reducing principal supervisors' span of control. It is difficult for supervisors to effectively support a large number of principals, especially when the role includes regular school visits with ongoing feedback to principals. Thus, reducing the span of control—or the number of

principals each supervisor oversees—is a central component of the PSI. Districts are expected to sustain the reduced spans of control with internal funding by the end of the initiative.

Figure I.2. The five core components of the PSI



Training supervisors and developing their capacity to support principals. The significant shift in the revised role for supervisors requires new and different areas of expertise than previously required for the position. To address these needs, districts work with external technical assistance providers to develop supervisors' capacity and skills to coach, mentor, and provide professional development for principals and to manage learning communities.

Developing systems to identify and train new supervisors (succession planning). Districts identify and develop new talent to fill future principal supervisor positions by creating a cadre of new supervisors with the requisite capacities for the revised supervisor role. Approaches to this component can include apprenticeship programs to prepare future principal supervisors to step into the position or developing leader tracking systems to identify and prepare prospective principal supervisors.

Strengthening central office structures to support and sustain changes in the principal supervisors' role. The change in principal supervisors' role requires shifting many managerial tasks that supervisors previously handled to other central office personnel; central office culture and structures also change to align with and support the new supervisors' role. Furthermore, as

their work shifts away from administration, compliance, and operations, supervisors work more closely with other academics-oriented departments, such as curriculum and instruction and teaching and learning. Although The Wallace Foundation asked that districts create a detailed plan for central office change only by the conclusion of the initiative, districts began to address this component in the early years of the initiative.

C. Looking back: A summary of the first three years of the PSI

As described in the initial report on the PSI's implementation (Goldring et al. 2018), districts made many changes in their approaches to principal supervision during the first three years of the initiative. However, each district determined how and when to address each component of the PSI. As a result, districts' emphases on particular components varied in the initial years of implementation, and they addressed the components at different times. In the first report, we documented how the six PSI districts demonstrated the feasibility of making substantial changes to the principal supervisor role, across all components of the initiative. The districts revised the job descriptions for principal supervisors, reduced the span of control, implemented new training programs, and restructured roles and responsibilities in the central office to support changes to the principal supervisor role. Some districts took steps to strengthen their succession planning.

These changes in the principal supervisor role laid the groundwork for changes in principal supervisors' day-to-day work with principals. By the end of the PSI's third year, most principal supervisors spent the largest share of their time in schools engaging in newly developed routines and practices, such as participating in classroom walk-throughs, coaching principals, and providing them with ongoing feedback. Principal supervisors also consistently met with groups of principals to provide opportunities for collaborative learning. They focused less on administration and building operations than in the past. They also focused less on compliance activities, such as ensuring district and state forms were completed correctly and submitted on time.

D. District context in the final year of the PSI

Each district's particular circumstances in the final year of the PSI influenced its experiences and progress with the initiative. One district faced budget shortfalls that required cuts to many central office departments and positions, along with the elimination of some principal supervisor positions. Some districts experienced leadership changes or central office reorganization, along with changing district priorities. Other districts, in contrast, benefited from longstanding stability in central office leadership and strong school board support. All districts faced challenges as they continued to focus on how best to meet the needs of low-performing schools in complex urban environments.

E. Goals and overview of this report

This report describes districts' experiences with the PSI across the years of the initiative and analyzes its effects on teachers' perceptions of principals' performance. It discusses districts' successes and challenges implementing the PSI and lessons learned for long-term sustainability. Specifically, after discussing the study methods in Chapter II, the report addresses the following questions:

- What is the theory of change for revising the principal supervisor's role? (Chapter III)
- What were districts' and principal supervisors' experiences with the PSI? (Chapter IV)
- What were principals' experiences with the PSI? (Chapter V)
- How did the initiative affect teachers' perceptions of principals' performance? How did teachers' perceptions of principals' performance and the effects of the PSI vary across districts and principal supervisors? (Chapter VI)
- How do key components of the PSI relate to changes in teachers' perceptions of principals' performance? (Chapter VII)
- How did the principal supervisor role in the PSI districts compare with that in other urban districts? (Chapter VIII)

The report's final chapter discusses the sustainability of changes made during the PSI and lessons learned for revising the principal supervisor role. The report also spotlights three topics: the role of the superintendent in the implementation of the PSI, principal supervisors' differentiation of support for principals, and supervisors' approaches to principal evaluation.

The report is based on analyses of data from multiple sources. We conducted semistructured interviews with the superintendent, and a sample of central office personnel, principal supervisors, and principals in each of the six PSI districts over the course of the initiative. Through the interviews, we aimed to understand the districts' ongoing experiences as they changed the principal supervisor role. We also administered surveys to all principal supervisors and principals in the PSI districts three times during the initiative. The survey data provided the perspectives of a broader range of supervisors and principals regarding each of the components of the PSI. We used data from the Vanderbilt Assessment of Leadership in Education (VAL-ED) from the 2013–2014 through 2017–2018 school years to analyze the PSI's effects on teachers' perceptions of principals' performance. To compare the principal supervisor role in PSI districts with that in other urban districts, we surveyed principal supervisors in a sample of urban districts that were not part of the PSI.

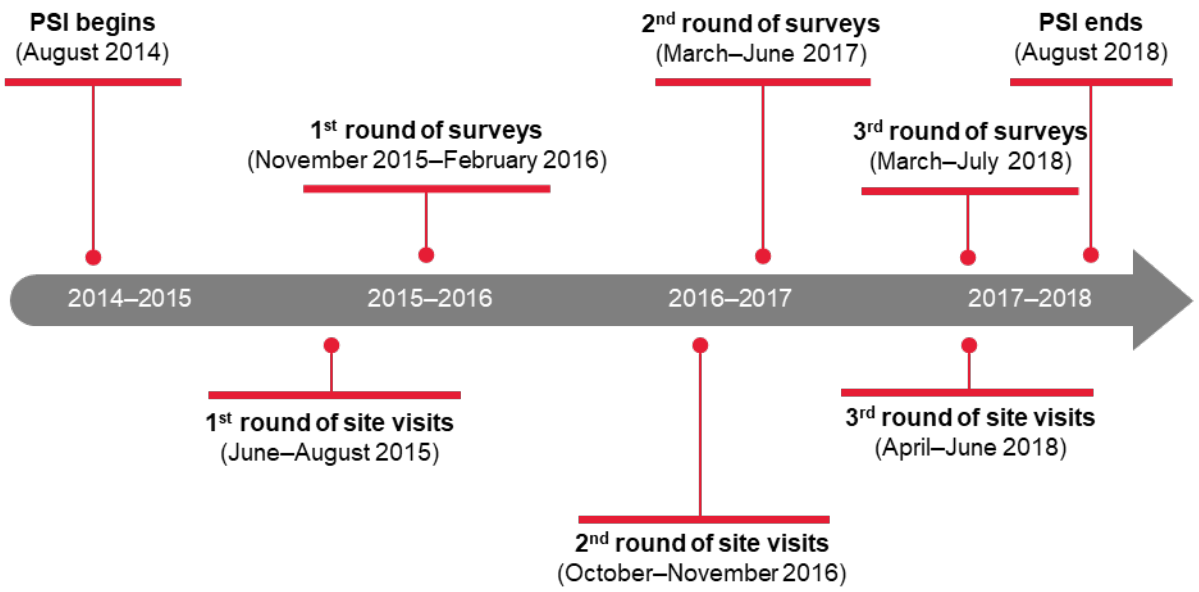
II. METHODS

This chapter describes the methods used to analyze (1) districts', supervisors', and principals' experiences with the PSI; (2) the PSI's effects on teachers' perceptions of principals' performance; (3) how teachers' perceptions of principals' performance and the PSI's effects vary across districts and supervisors; and (4) the relationship between districts' implementation of the PSI and its effects on teachers' perceptions of principals' performance. Appendix A provides additional details.

A. Examining districts', supervisors', and principals' experiences with the PSI

To study districts', supervisors', and principals' experiences with the PSI, we conducted site visits and collected surveys in the six PSI districts during the second, third, and fourth years of the initiative (the 2015–2016 through 2017–2018 school years) (Figure II.1). Data collection began after districts had already begun to implement the PSI.

Figure II.1 Timeline of data collection to measure experiences with the PSI



Site visits and interviews in PSI districts. The study team visited each of the six PSI districts three times: in summer 2015, fall 2016, and spring 2018.[1] The first visits took place about a year after the PSI began. Each visit lasted two to four days. During each visit, the study team conducted 18–20 semistructured interviews with central office personnel, principal supervisors, and principals. We developed the protocols for these interviews based on a review

[1] Because Baltimore did not join the initiative until fall 2015, the study team first visited that district in spring 2016.

of existing research on principal supervision and revised them for each subsequent round of data collection to include questions on emerging themes.

In each round of interviews, we focused on central office leaders', principal supervisors', and principals' understanding of the five components of the PSI, as well as their experiences related to these components. In the final year of the PSI, we also included additional interview questions on the following summative topics:

- Successes and challenges of the PSI
- Changes in the principal supervisor role
- Supervisors' differentiation of their practices to meet the diverse needs of principals and schools
- The role of the superintendent in the PSI
- Principal evaluation and principal supervision
- Districts' lessons learned from the PSI
- Districts' plans for sustaining changes made during the PSI

During visits, the study team also collected relevant artifacts and documents, such as agendas for principal supervisor trainings and principal supervisor job descriptions. We recorded and transcribed all interviews for analysis. In total, the study team conducted 335 interviews over three years across the six PSI districts (Table II.1).

Table II.1. Number and roles of interview participants

Role	Respondents per district in each site visit	Total in first site visit (summer 2015)	Total in second site visit (fall 2016)	Total in third site visit (spring 2018)	Total across all three site visits
Central office staff[a]	2–4	15	13	18[a]	46
Principal supervisor	6	35	37	40	112
Principal	10	59	60	58	177
Total	**18–20**	**109**	**110**	**116**	**335**

[a] In the final year of interviews, we included superintendents in the sample of central office staff.

- **Central office staff.** The team interviewed two to four central office staff members in each district during each visit. Typically, we interviewed the district's project director for the PSI (the main point of contact for the initiative) and the supervisor of the principal supervisors or someone else in the central office who was closely involved in the initiative. In the third round of site visits, we also interviewed the superintendent in each PSI district. Where relevant, we also interviewed other central office staff who worked closely with the principal supervisors.
- **Principal supervisors.** The team interviewed six principal supervisors per district in the first two rounds of interviews, and up to eight principal supervisors in the third round. In districts with more than six supervisors, we selected a sample of six to participate in the first

round of interviews. In these districts, we first grouped supervisors according to the type of schools supervised (for example, by geography or grade level) and then randomly sampled supervisors from these groups, to ensure that we included supervisors from all types of schools. For the second and third rounds of site visits, we re-interviewed previously sampled supervisors. If a previously sampled supervisor was no longer working as a supervisor in the district, we randomly sampled a new supervisor. Similarly, in the third round of interviews, when we expanded the sample to eight supervisors per district, we randomly selected two additional supervisors from the established strata.

- **Principals.** In each district, we grouped principals according to their supervisors and then randomly selected a sample of 10 principals from across these groups to ensure that as many supervisors as possible were represented in the principal sample. We excluded principals who were new to their role in the year of the interview or whose supervisors were not included in the PSI (for example, in some districts, alternative schools and their supervisors were not part of the PSI). In addition, we gathered information on the principals' school grade levels and school performance. We used this information to ensure that the sample of principals included a range of school grade and performance levels. In the second and third rounds of interviews, we re-interviewed five of the original principal participants to capture their perspectives over time. In addition, we selected a sample of five new principals according to the procedures described previously.

Surveys in PSI districts. We conducted three rounds of surveys in each district: fall 2015, spring 2017, and spring 2018. During each round, we sent online surveys to all principals and supervisors in the six PSI districts. We developed the surveys based on preliminary analysis of interview data, a review of current supervisor competency standards (Council of Chief State School Officers 2015), and prior research on principal supervisors, with a focus on the PSI's core components. Surveys measured principal supervisor engagement in key practices and behaviors related to the PSI, such as the number of times they met with principals in various settings, the percentage of time they spent focusing on various topics with principals, and how frequently they engaged in practices that were the focus of the PSI ("focal practices") when working with principals. Surveys also measured the quality of supervisor training, the quality of district central office support for supervisors and principals, and supervisors' and principals' assessment of their supervisor's effectiveness (see Appendix A for additional details).[2] In the second and third rounds of supervisor surveys, we also included questions about how supervisors differentiated their practices across two specific schools. We reviewed, piloted, and revised all survey items to ensure their face validity and clarity. Response rates for both principals and supervisors were very high for all three rounds of surveys across all districts (Table II.2).

[2] The principal and supervisor surveys are available online at https://www.wallacefoundation.org/knowledge-center/Pages/Changing-the-Principal-Supervisor-Role-to-Better-Support-Principals.aspx.

Table II.2. Principal and principal supervisor survey response rates

	Round 1 (fall 2015)		Round 2 (spring 2017)		Round 3 (spring 2018)	
	Respondents	Response rate (percentage)	Respondents	Response rate (percentage)	Respondents	Response rate (percentage)
Principal survey						
Baltimore	122	87	110	80	103	76
Broward	221	98	217	94	219	98
Cleveland	97	91	101	90	99	92
Des Moines	60	100	63	100	56	100
Long Beach	85	97	85	96	78	90
Minneapolis	59	92	64	97	64	98
Overall	**644**	**94**	**640**	**92**	**619**	**92**
Principal supervisor survey						
Baltimore	11	100	9	100	9	82
Broward	10	100	15	100	15	93
Cleveland	7	100	7	100	7	100
Des Moines	6	100	6	100	7	100
Long Beach	10	100	8	100	8	100
Minneapolis	6	100	6	100	4	100
Overall	**50**	**100**	**51**	**100**	**50**	**96**

Analyses. The research team coded the interview data using an iterative process, focusing on each component of the PSI as well as on how each component unfolded and was experienced in each district by central office staff, supervisors and principals. The data analyses focused on districts', supervisors', and principals' experiences, successes, and challenges throughout the initiative; lessons learned; and plans for sustaining the accomplishments of the PSI after the initiative ended. To understand how supervisors differentiated their work with principals, we examined how supervisors varied their practices for two specific principals—one more experienced and one less experienced—based on data from the principal supervisor surveys. We integrate the descriptive survey results with the qualitative data. We report unweighted survey responses, giving equal weight to the perceptions and experiences of respondents across the initiative rather than weighting each district equally.

B. Estimating the PSI's effects on teachers' perceptions of principals' performance

To provide rigorous evidence of the effects of the PSI on teachers' perceptions of principals' performance, we identified a comparison group that was similar to the schools and principals in the PSI districts at the start of the study. We then tracked and compared the performance of principals in the two groups over time to estimate the effect of the PSI on teachers' perceptions of principals' performance. We describe our approach further below.

1. Measuring teachers' perceptions of principals' performance

We measured teachers' perceptions of principals' performance using teachers' ratings of principals from the VAL-ED.[3] The VAL-ED is a principal evaluation tool that combines feedback from surveys of principal supervisors, principals, and teachers to provide an overall measure of principals' performance (Porter et al. 2010a, 2010b). It is well-suited to measuring the effects of the PSI because it focuses on the same types of instructional leadership practices and competencies emphasized by the PSI. Principal supervisors, principals, and teachers all rate principals on the VAL-ED instrument, which contains 72 items, each on a 5-point scale.[4]

The VAL-ED focuses on six core components of school performance and six key processes of effective school leaders:

Core components	Key processes
1. High standards for student learning	1. Planning
2. Rigorous curriculum	2. Implementing
3. Quality instruction	3. Supporting
4. Culture of learning and professional behavior	4. Advocating
5. Connections to external communities	5. Communicating
6. Performance accountability	6. Monitoring

(Porter et al. 2008)

Although districts typically focus on the aggregate VAL-ED scores that combine feedback from principal supervisors, principals, and teachers, we focused on teachers' ratings due to concerns that the aggregate measure could be biased. Such bias might arise because the PSI might have affected not only the true performance of principals, but also *how* supervisors rated principals and how principals rated themselves. For instance, the initiative could have led principal supervisors to hold principals to a higher standard, and principals to hold themselves to a higher standard. If so, it might not be appropriate to compare aggregate VAL-ED scores in the PSI schools with those in comparison schools. However, because teachers were not directly involved in the PSI, the way they rated their principals is unlikely to have been affected by the initiative. Teacher, principal, and principal supervisor ratings are positively correlated,

[3] The Wallace Foundation funded the development of the VAL-ED, which was coauthored by Ellen Goldring and colleagues and copyrighted by Vanderbilt University. Mathematica collected and analyzed the VAL-ED data for this study, and The Wallace Foundation assisted with data collection. Dr. Goldring was not directly involved in the collection or analysis of VAL-ED data for this study.

[4] Principal supervisors and principals complete the full assessment; teachers are randomly assigned to complete one of two halves to reduce response time with little reduction in reliability (Porter et al. 2008).

suggesting that the teacher scores capture a similar underlying construct as those from other raters (Appendix A, Table A.8).

The Wallace Foundation required all PSI districts to administer the VAL-ED in spring 2014, as part of the grant application process, and for each subsequent year of the initiative.[5] The foundation encouraged districts to attain a teacher response rate of at least 75 percent in each school. We analyzed VAL-ED scores from 2013–2014 (the baseline school year) and all subsequent years through 2017–2018. We compared scores between schools in the PSI districts and a set of comparison schools in other districts that also administered the VAL-ED in the relevant years, described further below.

2. Constructing a matched comparison group

To learn how the PSI affected teachers' perceptions of principals' performance in the PSI schools, we constructed a matched comparison group of schools in non-PSI districts that were similar to the sample of PSI schools before the start of the initiative. We conducted the analysis at the school level rather than at the principal level. This approach allowed us to estimate the full effect of the PSI on teachers' perceptions of principals' performance in the PSI schools, taking into account that some of the PSI's effect could be driven by changes in principal turnover rather than changes in the performance of a fixed set of principals. For example, if the PSI induced districts to dismiss more low-performing principals and retain more high-performing ones, an analysis focusing on changes in teachers' perceptions of principals' performance in a fixed set of PSI schools (rather than on changes in the performance of a fixed set of principals) would capture these improvements in teachers' perceptions of principals' performance due to effects on principal turnover as part of the overall PSI effect.

The sample of PSI schools included all schools with VAL-ED scores that participated in the PSI for the full implementation period in their districts.[6] Almost all schools in each PSI district were included in the initiative from the start. The sample includes teacher-reported VAL-ED scores for more than 99 percent of the principals in the PSI. However, in some PSI districts, some additional schools that were not part of the PSI also administered the VAL-ED. For example, in one district, alternative schools administered the VAL-ED even though they did not have principal supervisors who participated in the PSI and therefore were not part of the initiative. We excluded any schools in the PSI districts that were not part of the initiative. We also excluded two charter schools that were taken over by a PSI district midway through the initiative so that we could capture the effects of participation in the PSI for the full length of the initiative.

To identify the matched comparison group, we began with a set of *potential* comparison schools from non-PSI districts that administered the VAL-ED between 2013–2014 and 2017–2018. We further limited the sample to urban and suburban schools because the schools in all six

[5] Even though Baltimore joined the PSI late (in fall 2015), it administered the VAL-ED in both spring 2014 and spring 2015.

[6] In all districts but Baltimore, implementation began in fall 2014. Baltimore joined the initiative late and began implementation in fall 2015. For this reason, we exclude Baltimore from the 2014–2015 analysis sample but include it in all later years. Implementation in all six districts continued through spring 2018.

PSI districts fell into one of these two categories. We also excluded districts participating in The Wallace Foundation's Principal Pipeline Initiative and two districts that received smaller PSI grants from the foundation as PSI exemplar districts because these districts received similar supports to the six PSI districts in the study.

We then used an approach known as propensity score matching to identify comparison schools that were similar to the PSI schools before the initiative began. In particular, for each PSI school, we identified a set of comparison schools that were similar on several key dimensions that might be correlated with both district participation in the PSI and principals' effectiveness. These included baseline VAL-ED scores, school demographics, and student achievement. Table II.3 shows the full set of characteristics used to form comparison schools.

Table II.3. Baseline characteristics of PSI and matched comparison schools (2017–2018 analysis sample)

Variable	PSI school mean	Matched comparison school mean	Difference between PSI and matched comparison school means
School characteristics			
School size, level, and locale			
Total enrollment	736	700	36
Number of full-time equivalent teachers	41	40	1
Percentage elementary school	46	27	19
Percentage high school	54	73	–19
Percentage urban locale	69	69	0
Percentage suburban locale	31	31	0
School demographic composition			
Percentage black, non-Hispanic	49	56	–7
Percentage white, non-Hispanic	20	17	3
Percentage Hispanic	23	21	2
Percentage other race	8	6	2
Percentage eligible for free or reduced-price lunch	74	78	–4
Student achievement			
School math and reading proficiency rates			
Math, baseline year	51	49	2
Math, one year prior	55	52	3
Math, two years prior	57	53	4
Reading, baseline year	56	55	1
Reading, one year prior	56	56	0
Reading, two years prior	58	56	2
VAL-ED scores			
VAL-ED scores (teacher's perceptions of principals' performance)	3.6	3.5	0.1
Percentage of teachers responding to the VAL-ED survey	80	81	–1

Sources: VAL-ED survey data (2013–2014 school year), Common Core of Data (2013–2014 school year), and EdFacts data (2011–2012 to 2013–2014 school years).

Note: All variables are from the baseline (2013–2014) school year unless otherwise indicated. The mean of the comparison schools was calculated using the weights derived from the propensity score matching procedure. The sample includes all schools that enter the analysis during the 2017–2018 school year. The sample excludes districts that are in The Wallace Foundation's Principal Pipeline Initiative and two exemplar districts that received smaller PSI grants but are not part of the evaluation. Because all schools in PSI districts were classified as either urban or suburban, we also restrict the comparison schools to those classified as urban or suburban. None of the differences between the PSI and matched comparison schools were statistically significant at the 5 percent level.

The propensity score matching approach resulted in a well-matched sample of PSI and comparison schools. Table II.3 shows that the PSI and matched comparison schools had similar values on key baseline variables for the 2017–2018 analysis sample. For example, the PSI and comparison schools differed in baseline VAL-ED scores by less than 0.10 out of a possible 5 points. None of the differences between the PSI and matched comparison schools were statistically significant at the 5 percent level. We also conducted formal tests for the analysis sample for each year, in which we calculated the differences in the baseline characteristics between the PSI and matched comparison schools (see Appendix A for a full description). Of the 64 estimated differences, only two exceeded 0.25 standard deviations (see Appendix Table A.10), a standard threshold for whether a difference is large (Stuart 2010). These two exceptions were limited to the 2016–2017 and 2017–2018 school years, when fewer districts administered the VAL-ED and sample sizes were smaller. With smaller sample sizes, it is more likely to estimate a large difference due to chance. Supporting this possibility, neither of these two differences was statistically significantly different from zero.

3. Estimating of the PSI's effects on teachers' perceptions of principals' performance

To estimate the overall effect of the PSI for all schools in the initiative, we first estimated the effect for each PSI school individually. To do so, we calculated the difference between the VAL-ED score for each PSI school and the average VAL-ED score for that school's matched comparison schools. We then averaged these school-level effects to estimate the overall effect of the PSI. We also used the estimated school-level effects for the analysis of variation in PSI's effects described in Section C and the correlational analyses described in Section D.

In addition, we examined how the PSI's effects differed between elementary and secondary schools. Secondary schools tend to be larger than elementary schools, so principals might have less direct influence on teachers, curriculum, and instruction. This could influence the role that principal supervisors play in supporting secondary school principals, as well as the effects of the PSI. To explore how the PSI's effects differed for primary and secondary schools, we used a similar approach to our main analysis, estimating separate models by school level. We defined elementary schools as those for which all students are in grades six or lower and secondary schools as all other schools.

4. Statistical power

The sample sizes for the analysis provided sufficient statistical power to detect meaningful effects in all four years of the PSI (Table II.4). Between the 2015–2016 and 2016–2017 school years, the number of potential comparison schools and districts fell, likely because the VAL-ED license was transferred to a new licensee during this time period, and many districts did not renew their contracts. Nevertheless, these sample sizes provided adequate power to detect substantively meaningful effects. In this context, we consider a substantively meaningful effect to be equivalent to changing an average principal into a high-performing principal. We defined average as scoring in the 50th percentile on the baseline VAL-ED among all principals with VAL-ED scores in the 2013–2014 school year and high-performing as scoring in the 75th percentile. Under this definition, a meaningful effect is equivalent to a 0.30-point increase in

VAL-ED scores (or, for example, an increase from a VAL-ED score of 3.70 to a score of 4.00).[7] Our main analysis was large enough to detect a meaningful effect in all years.

Table II.4. Sample sizes and estimated minimum detectable effect sizes

Sample/units	2014–2015	2015–2016	2016–2017	2017–2018
Sample sizes				
PSI districts	5	6	6	6
PSI schools	490	628	585	612
Matched comparison districts	91	69	22	20
Matched comparison schools	327	347	92	208
Minimum detectable effects				
VAL-ED score units	0.25	0.17	0.30	0.29
Standard deviation units	0.52	0.33	0.62	0.57

Sources: VAL-ED survey data (2013–2014 to 2017–2018 school years), Common Core of Data (2013–2014 school year), and EdFacts data (2011–2012 to 2013–2014 school years).

Note: The counts refer to the number of schools and districts that enter the analysis for each year. The number of matched comparison schools refers to the number of comparison schools that serve as a comparison for at least one PSI school. The sample excludes districts that are in The Wallace Foundation's Principal Pipeline Initiative and two exemplar districts that received smaller PSI grants but are not part of the evaluation. Because all schools in PSI districts were classified as either urban or suburban, we also restrict the comparison schools to those classified as urban or suburban. The sample excludes Baltimore during the 2014–2015 school year because Baltimore had not yet implemented the PSI during this time. The minimum detectable effects in VAL-ED score units and standard deviation units are based on the following formulae (respectively): $MDE = \left[T_{df}^{-1}\left(1-\frac{\alpha}{2}\right) + T_{df}^{-1}(\beta)\right] * SE$ and $MDES = \left[T_{df}^{-1}\left(1-\frac{\alpha}{2}\right) + T_{df}^{-1}(\beta)\right] * SE/\sigma_Y$, where T_{df}^{-1} is the inverse student's t distribution with df degrees of freedom, α is the significance level set to be 0.05, β is the power set to be 0.80, SE is the standard error, and σ_Y is the standard deviation of the VAL-ED score. The df is the number of districts in the sample minus 1.

5. Limitations of approach

Although we believe this approach to estimating the PSI's effects provides the most rigorous possible estimates, it has some limitations.

First, we were only able to examine one outcome measure: teachers' perceptions of principals' performance as measured by the VAL-ED. Although the VAL-ED is a valid and reliable measure of teachers' perceptions of principals' performance, it may not capture all dimensions of principal's performance. The PSI could have affected other aspects of performance not captured by VAL-ED. It could also have affected other outcomes of interest to districts, including principal retention or satisfaction. We were not able to measure these outcomes in both PSI and comparison schools, so we could not include them in our analysis of the PSI's effects.

Second, we cannot draw definitive conclusions about whether the PSI caused any changes we find in teachers' perceptions of principals' performance, and we must therefore interpret the estimated effects with caution. Our matching procedure matched schools in the PSI districts to similar schools in non-PSI districts. Although we were able to match schools on a rich set of

[7] An increase of .30 points on the VAL-ED is equal to an effect size (the increase divided by the standard deviation of VAL-ED scores) of 0.62.

variables that are likely to be correlated with changes in teachers' perceptions of principals' performance and district participation in the PSI, there could be other characteristics we did not observe in our data that were correlated with changes in teachers' perceptions of principals' performance. To the extent that such characteristics were excluded from our matching models, the matched comparison groups might not have provided an accurate estimate of how principals in the PSI districts *would have performed* had they not been in a PSI district.

C. Examining how teachers' perceptions of principals' performance and the PSI's effects vary across districts and supervisors

In addition to examining the PSI's average effects on teachers' perceptions of principals' performance, we also examined how teachers' perceptions of principals' performance and the PSI's effects varied across schools and districts.

Examining how teachers' perceptions of principals' performance varied across schools and districts can contribute to our understanding of the extent to which principal supervisors and district policies influence principals' performance and in turn provide context for our interpretation of effects. For example, if the performance levels of principals assigned to the same supervisor are no more similar to one another than to those of principals assigned to other supervisors, this could suggest that individual supervisors play a limited role in their principals' performance. Thus, interventions focused solely on principal supervisors might have limited effects on principals' performance compared with interventions that more directly interact with principals.

Similarly, understanding how the PSI's effects varied across schools and districts can help shed light on the role that principal supervisors and district policies play in the PSI's effects on teachers' perceptions of principals' performance. For example, if the PSI's effects were more similar within schools assigned to the same supervisor than across schools assigned to different supervisors, this could suggest that individual supervisors are an important determinant of the PSI's effects in the schools they oversee. Similarly, if the PSI's effects were more similar for schools within the same district, this could suggest that district-level implementation factors were an important determinant of the PSI's effects.

To explore how teachers' perceptions of principals' performance and the PSI's effects varied across schools and districts, we estimated a measure known as the intracluster correlation (ICC). The ICC measures the percentage of the overall variation in either outcome (teachers' perceptions of principals' performance or school-level PSI effects) that is accounted for by differences between supervisors and between districts. A finding that differences between supervisors (or districts) account for a large percentage of the variation would suggest a large influence of supervisors (or district factors) on teachers' perceptions of principals' performance or the PSI's effects. We measured teachers' perceptions of principals' performance using teacher-reported scores from the VAL-ED (the outcome variable used for the effects study), and we measured the PSI's effects using the school-level estimates of the PSI's effects on teachers' perceptions of principals' performance described earlier. Our analysis incorporates all years of available data (from 2014–2015 to 2017–2018 for variation across districts and from 2015–2016 to 2017–2018 for variation across supervisors). We provide a full description of our approach to this analysis, including our adjustments for measurement error, in Appendix A.

We also investigated whether variation in teachers' perceptions of principals' performance was similar outside of the PSI districts. To do so, we calculated district ICCs for teachers' perceptions of principals' performance in the matched comparison schools from the effects study. We present these results in Appendix B. (In the comparison schools, we did not have data on principal supervisors or estimates of the PSI's effects.)

D. Examining factors associated with the PSI's effects

To further explore the PSI's effects on teachers' perceptions of principals' performance, we examined how three types of implementation factors related to the PSI's effects in each school: (1) implementation of specific components of the PSI; (2) principal supervisors' time spent on instructional leadership; and (3) supervisors' effectiveness, as perceived by principals (Table II.5). To measure these factors, we constructed scales using data from the principal and principal supervisor surveys. Appendix A provides more details on the construction of these scales.

Table II.5. Implementation measures

Specific components of the PSI
Supervisor's span of control (number of principals overseen) (single item)
Supervisor training quality (scale)
Alignment of supervisor training with PSI goals (scale)
Quality of central office support according to supervisors (scale)
Quality of central office support according to principals (scale)
Principal supervisors' time spent on instructional leadership
Supervisor's total time spent with principals on instructional leadership in past 3 months, as reported by principals (minutes) (composite measure)
Percentage of supervisor's total time spent with principals that was spent on instructional leadership activities in past 3 months, as reported by principals (single item)
Supervisors' effectiveness, as perceived by principals
Principals' ratings of supervisor's effectiveness (scale)
Supervisor's implementation of PSI's focal practices (scale)
Principals' perceptions of evaluation feedback (scale)

Note: We examined only measures of implementation that we expected to vary within districts. We did not construct measures of the two PSI components that did not vary within districts: revising the principal supervisors' job description to focus on instructional leadership, and developing systems to identify and train new supervisors (succession planning).

To investigate how each of these implementation factors related to the PSI's effects on teachers' perceptions of principal's performance, we modeled the relationship between PSI effects for each school and each of the implementation measures. In our main models, we used school effects from the final year of the initiative, reflecting the PSI's cumulative effects on teachers' perceptions of principals' performance over four years of the initiative. However, because implementation across all years of the initiative may have contributed to these cumulative effects, we averaged the implementation measures across the 2015–2016, 2016–2017, and 2017–2018 school years (the three years of the PSI during which data were collected). We also conducted sensitivity tests to examine whether findings were robust to other models and samples, including relating implementation factors in each year to the PSI's cumulative effects as of that year. (Appendix A provides more details on the regression model and sensitivity tests.)

This analysis can suggest hypotheses about features of PSI implementation that may have influenced its effects. However, it is an exploratory analysis and less rigorous than the main

analysis of PSI effects. In particular, it can suggest but not confirm causal relationships between measures of PSI implementation and effects.

E. Comparing the principal supervisor role in PSI districts and other urban districts

To provide context for the changes to principal supervision in the PSI districts, we compared the experiences of principal supervisors in the PSI districts with those of principal supervisors in other urban districts. To do so, we administered a survey to principal supervisors in urban districts that were members of the Council of the Great City Schools but were not part of the PSI. The survey included a subset of the questions asked on the principal supervisor survey administered in the PSI districts, covering topics such as supervisors' span of control, time use and focus, and perceptions of the quality of their districts' principal evaluation system and central office. We administered the survey in spring 2018, at the same time we surveyed principal supervisors in the PSI districts.

We administered the survey to all principal supervisors in 57 urban districts that were members of the Council of the Great City Schools.[8] In four of these districts, no supervisors responded, leaving us with a sample of 53 districts. We also excluded five other districts from the analysis sample, leaving us with a final analysis sample of 48 districts. Two of the excluded districts—District of Columbia Public Schools and Tulsa Public Schools—were engaged in separate but concurrent principal supervision changes with support from The Wallace Foundation. The other three—Charlotte-Mecklenburg Schools, Denver Public Schools, and Hillsborough County Public Schools—participated in the Principal Pipeline Initiative (PPI), a separate Wallace Foundation effort that also had implications for the principal supervisor role (the other three PPI districts were not members of the Council of the Great City Schools). Excluding these districts allowed us to compare supervisors in PSI districts with supervisors in urban districts that were not participating in intensive, targeted work that included some of the PSI's approaches and components. Although the other urban districts in our sample may also have made changes to their principal supervisor role during the same period, these districts likely implemented such changes on their own, without external support.

A total of 293 supervisors from the final set of 48 districts responded to the survey (a 64 percent response rate). When combined with the sample of 50 principal supervisors from the survey administered in the six PSI districts, the final sample for the analysis comparing PSI districts with other urban districts included 343 confirmed principal supervisors across 54 districts (a 67 percent response rate).

In the analysis, we compared the survey responses of principal supervisors in PSI districts and other urban districts. As in the analysis of PSI districts alone, we report unweighted survey responses, giving equal weight to the perceptions and experiences of respondents across all the districts in the sample rather than weighting each district equally.

[8] The Council of the Great City Schools had 70 member districts at the time of survey administration in 2018, including the 6 PSI districts and 64 non-PSI districts. Of the 64 non-PSI districts that were members of the Council, 7 did not provide a list of their principal supervisors to the study team and were not included in the survey, leading to the sample of 57 non-PSI districts.

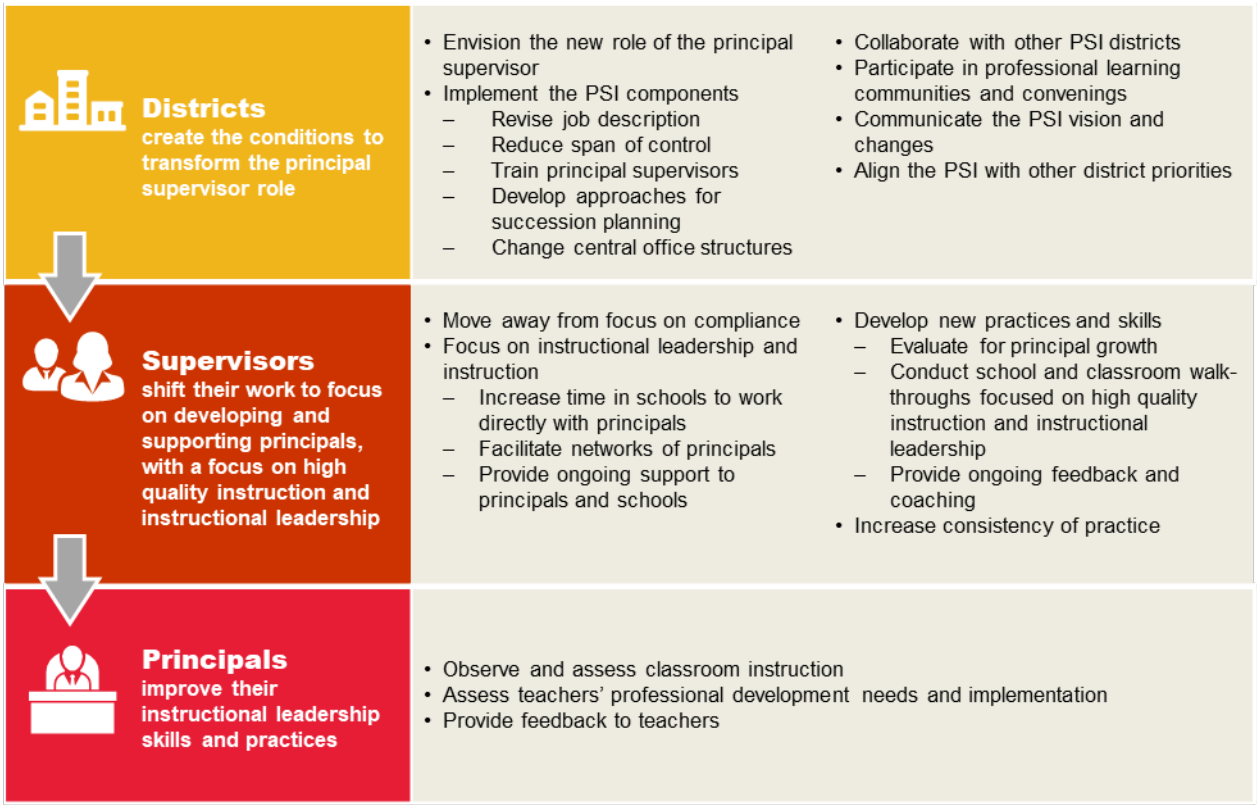
III. THE THEORY OF CHANGE FOR REVISING THE PRINCIPAL SUPERVISOR ROLE

The PSI was a district-level initiative that sought to improve principals' effectiveness by changing the principal supervisor role. Specifically, the five components of the initiative included changes to central office structures and policies to better support supervisors in their work with principals. Notably, the initiative did not directly touch principals—for example, it did not provide direct training for principals or ask districts to create pipeline programs to develop new principals. Rather, the PSI's five components were intended to create the conditions necessary for districts and supervisors to support and develop principals.

Previous research has shown that the knowledge and skills of central office staff directly influence school districts' capacity to implement instructional improvement strategies at the school level (Burch and Spillane 2004; Marsh et al. 2005). Principal supervisors, as the central office personnel directly responsible for overseeing principals' development, benefit from structures, policies, and training that deepen their knowledge of teaching and learning and facilitate their ability to work with principals as coaches and teachers rather than as compliance managers (Honig et al. 2010). The PSI provides an opportunity to learn whether a district-level intervention that does not directly interact with principals can improve principals' performance.

Although the PSI laid out a broad approach to overhauling the principal supervisor role through its five core components, it did not directly define the specific approaches districts should take to achieve this goal, nor did it specify particular skills or practices for principal supervisors. Rather, it encouraged participating districts to implement the components in ways that aligned with their particular contexts and priorities. The order in which districts addressed the components was also very flexible. Therefore, although the PSI did not lay out a specific theory of change for revising the principal supervisor role, this theory emerged over the course of the initiative as districts adapted the PSI components to define, change, and support the supervisor role according to their unique goals and contexts. Figure III.1 describes the theory of change that emerged from the PSI districts over the four years of the initiative, based on the field research, as districts sought to change the principal supervisor role and improve principals' effectiveness.

Figure III.1. The theory of change that emerged as districts revised the principal supervisor role



District-level actions. The PSI focused districts on several specific actions to create the conditions to transform the principal supervisor role:

- Defining the responsibilities of the position and its core practices through revising the job description, which was one of the five core components of the PSI.
- Deliberating and debating about how supervisors should work and interact with principals.
- Implementing the other components of the PSI, including reducing the span of control, implementing dedicated programs to train supervisors for their new roles, developing approaches to succession planning to fill new supervisor positions, and restructuring the central office to better support and coordinate with supervisors.
- Aligning the PSI with other district priorities.
- Communicating the district's vision for the PSI and associated changes to various stakeholders, such as school boards, central office personnel, and principals.
- Shifting noninstructional responsibilities, such as maintenance, away from supervisors so that they could focus on supporting principals' instructional leadership development.
- Changing the central office to be more responsive to schools.

All these changes were intended to increase supervisors' focus on instructional leadership through increased one-on-one interactions and meetings in smaller networks of principals. The PSI also provided networking supports for districts through a series of professional learning community meetings convened by The Wallace Foundation. These meetings facilitated communication and collaboration among the six PSI districts, national organizations, and districts participating in other Wallace Foundation initiatives.

Supervisor-level actions. At the supervisor level, the PSI defined two actions for changing the principal supervisor role: (1) move away from focusing on compliance with principals and (2) focus on instructional leadership and instructional quality in their work with principals. Beyond setting out these goals, the PSI did not define specific approaches supervisors should take to accomplish them.

Over the course of the PSI, the district-level actions helped to develop and support principal supervisors in their efforts to improve principals' instructional leadership skills and practices. Although the PSI did not explicitly define specific changes to the principal supervisor role, some common practices and approaches emerged across the six districts. These approaches were designed to shift the principal supervisor role away from compliance and administrative tasks toward developing and supporting principals. As supervisors developed a core set of practices and acquired skills to implement them, this was expected to improve consistency from one supervisor to the next, ensuring more consistent supports and evaluation for all principals. Supervisors would spend more time in schools as districts removed administrative responsibilities they held in the past; supervisors would have deeper and more specific understanding of the needs of principals and their schools and would provide targeted supports as districts reduced their span of control. These changes in the supervisor role would increase supervisors' total time spent on supporting principals, facilitating networks of principals, and creating consistency of practice among supervisors in each district.

Principal supervisor training was intended to develop supervisors' skills to implement new practices, such as:

- conducting classroom walk-throughs to assess instructional quality and rigor
- evaluating and assessing principals' growth in instructional leadership skills
- providing actionable feedback and ongoing coaching to principals.

Principal-level actions. The PSI did not prescribe any specific principal-level actions. The components of the PSI were all geared toward actions that districts initiated with their principal supervisors. Districts that agreed to participate in the PSI were asked not to implement other specific principal focused initiatives, such as pipeline programs, to keep the focus on the PSI. Instead, supervisors worked with and coached the principals in their networks to develop principals' instructional leadership skills, such as observing and assessing classroom instruction, helping teachers implement high-quality instructional practices, and providing actionable teacher feedback. The PSI surmised that supervisors, with their reduced span of control, central office support, and enhanced skills, would be able to better develop principals' instructional leadership practices. Through ongoing coaching and support from principal supervisors, principals would enhance their ability to observe classrooms and provide feedback to teachers, identify

professional development needs of staff, and design and facilitate their professional development to strengthen the learning environment for students.

Summary. The PSI focused on district-level changes to redefine the principal supervisor role to support and improve principals' performance. The initiative gave wide leeway to districts in how they implemented the changes and components of the PSI. Districts varied in the specific approaches and structures they created for and around the supervisor role. Over the course of the initiative, we learned that the task of changing the principal supervisor role was closely linked to broader changes within the districts' central office, as discussed further in Chapter IV.

IV. DISTRICTS' AND PRINCIPAL SUPERVISORS' EXPERIENCES WITH THE PSI

As the PSI came to a close, principal supervisors and district leaders reflected on the changes made in their districts during the initiative. Consistent with the PSI's theory of change, the districts made significant changes to the principal supervisor role and to the central office to better support supervisors' work. Districts' and principal supervisors' experiences with the PSI highlight both the challenges and the rewards of undertaking district-level changes.

A. Districts' experiences

The revised principal supervisor job description guided districts' changes to the principal supervisor role. All districts revised their principal supervisor job descriptions as an early step in the PSI. These job descriptions served as a roadmap to guide other changes over the course of the initiative. For example, a district leader in Des Moines explained how the job description guided the training that the district provided to supervisors.

> **We just have deepened [supervisors'] knowledge, because when we looked at their job description, they are supposed to support principals around curriculum and instruction, but then we realized we never really provided them with professional development or support.**
>
> ***—District leader (2018)***

Districts reduced principal supervisors' span of control over the course of the initiative. Before the PSI, supervisors across the six districts oversaw an average of 17 principals. The initiative suggested that districts target a span of control of 12 principals per supervisor. Each district reduced supervisors' span of control early in the initiative, with four of the six districts reducing span of control in the first two years. By the end of the initiative, the average span of control across all six districts had fallen to 13 principals (Table IV.1).

- Des Moines added a new supervisor position in the PSI's final year, which allowed it to add a new network (one supervisor and the group of principals he or she supervised) that served all the International Baccalaureate schools in the district; these schools were previously included in other networks.
- Broward was planning to add one additional position for the 2018–2019 school year with hopes of reducing span of control at the secondary level.
- Cleveland moved some of the lowest-performing schools into a single "portfolio network" that operated separately from the main networks. This reduced the number of schools that the other supervisors oversaw.
- Baltimore was able to maintain its reduced span of control throughout the initiative.
- Minneapolis increased the span of control during the PSI.

Table IV.1. Span of control over the course of the PSI

	Pre-PSI	2015–2016		2016–2017		2017–2018	
	Mean	Mean	Range	Mean	Range	Mean	Range
Baltimore	13	13	(9–16)	14	(12–17)	13	(10–16)
Broward	21	21	(15–25)	15	(11–19)	14	(11–19)
Cleveland	16	14	(14–15)	13	(11–15)	12	(9–15)
Des Moines	16	10	(8–11)	10	(8–11)	8	(7–10)
Long Beach	17	9	(3–12)	11	(9–13)	11	(9–13)
Minneapolis	17	10	(7–12)	10	(6–13)	16	(14–17)
Overall	**17**	**13**	**(3–25)**	**12**	**(6–19)**	**13**	**(7–19)**

Sources: The Wallace Foundation 2014 (Pre-PSI) and Principal supervisor surveys (2015–2018).

Note: Span of control is defined as the number of principals assigned to and overseen by a single supervisor. It does not refer to teams in which multiple staff share principal supervision responsibilities (there were no such teams in the PSI districts). The overall mean is calculated as the mean span of control across all supervisors.

Principal supervisors with smaller spans of control visited principals more frequently. When principal supervisors oversaw fewer principals, their principals reported having more frequent meetings with them at their schools. Figure IV.1 displays the average number of principal-reported meetings with their supervisor at their school across the range of supervisor spans of control. Supervisors with smaller spans of control (3–11 principals) visited their principals 5.2 times on average over a three-month period, compared with supervisors with larger spans of control (15–25 principals), who visited their principals an average of only 3.1 times. Overseeing more schools also meant spending more time in the car. Supervisors who oversaw 15 principals or more reported spending more time (289 minutes) traveling among district sites in a typical week compared with supervisors whose span of control was 11 or fewer principals (182 minutes) or 12–14 principals (206 minutes; not shown in figure).

Figure IV.1. Principals reported meeting with their supervisor at school more often when their supervisor's span of control was smaller

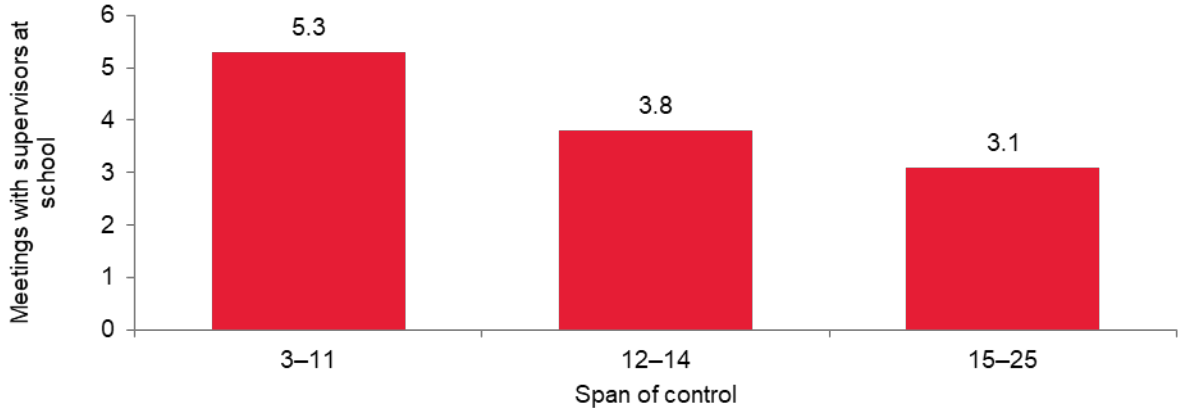


Figure reads: Principals whose supervisor's span of control was 3–11 principals reported meeting with their supervisor at school an average of 5.3 times over a three-month period.

Source: Principal surveys, 2016 (*N* = 635), 2017 (*N* = 639), and 2018 (*N* = 606).

Note: The survey question read, "Over the past three months, how many times have you met with your principal supervisor in the following settings, and how much time do you spend with him or her in a typical meeting in your school?"

Reduced span of control helped principal supervisors provide more intensive support to principals. Supervisors with smaller spans of control also placed more emphasis on key principal instructional leadership practices, including data use practices (for example, helping the principal analyze and interpret their school's data), feedback practices (for example, modeling or role-playing teacher feedback), and classroom visit practices (for example, working with the principal to give feedback based on classroom visits) (Figure IV.2).

Figure IV.2. Supervisors with larger spans of control placed less emphasis on key practices in 2018

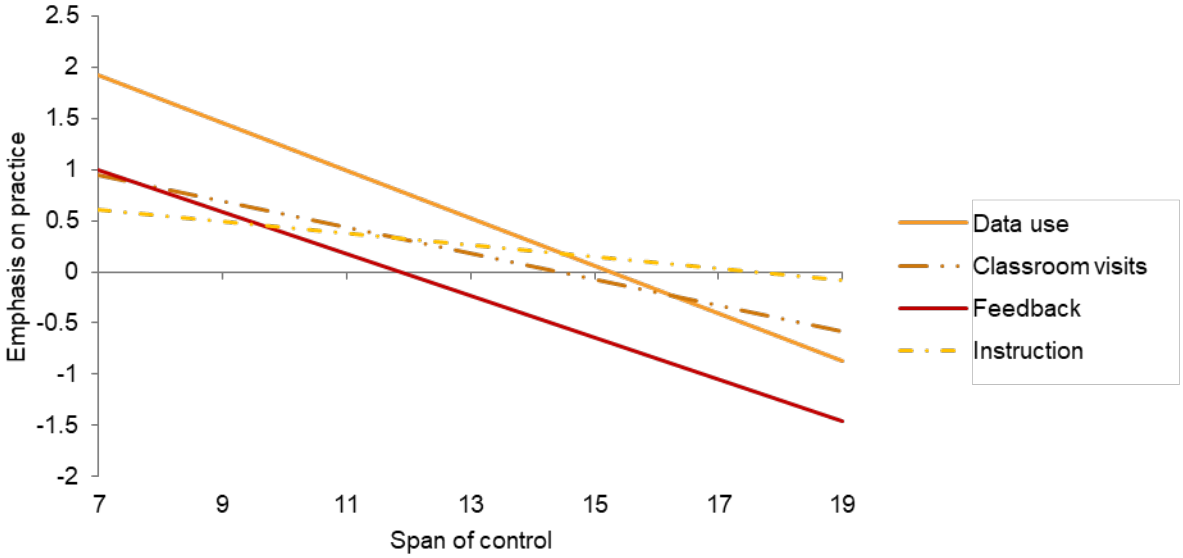


Figure reads: As supervisors' span of control increases, their emphasis on key instructional leadership practices with principals declines.

Source: Principal supervisor survey 2018 (*N* = 50).

Notes: Results are a fitted line between predicted outcomes and reported span of control from a regression analysis. Each outcome has been standardized to a mean of 0 and standard deviation of 1 (see Appendix A). The relationships with span of control are statistically significant at the 5 percent level for data use, classroom visits, and feedback.

Findings are based on supervisors' reports regarding the emphasis they placed on key practices of instructional leadership support with their least and most experienced principals.

Findings are based on supervisors' reports regarding the emphasis they placed on key practices of instructional leadership support with their least and most experienced principals. In interviews, supervisors with smaller spans of control said they were able to interact with all their principals more frequently, establish systems for monitoring principals' work more consistently, and clarify their expectations for principals.

> **[Reducing the span of control] allowed us to set up a system where we genuinely can have accountability, and that sounds so elementary, but that's a huge paradigm shift for our folks. Our principals are feeling it, teachers are feeling it, but at the same time, they're getting way more support than they've ever had before, so there's a lot more predictability for our leaders in this system. There's a lot more clarity about expectations … and a principal job at one building looks a lot more like a principal job at another building now.**
>
> ***—District leader (2018)***

Districts clarified their definitions of high-quality instruction and instructional leadership. At the beginning of the initiative, not all districts had clear definitions and common understandings of high-quality instruction and instructional leadership. By the end of the PSI, there were common definitions of high-quality instruction among supervisors and other central office personnel in all districts. Supervisors received training on the new instructional frameworks developed during the PSI and worked with principals to implement these frameworks in their schools.

- In Cleveland, supervisors received training on identifying and cultivating high-quality instruction. The training was developed and initially facilitated by an external technical assistance provider as a community of practice for supervisors. In the PSI's final year, the district hired a leadership development specialist to continue the training. Supervisors focused on helping principals evaluate whether instruction was high-quality, rigorous, and appropriate.
- In Des Moines, supervisors developed a common understanding of academic rigor and used this understanding to help principals in their classroom walk-throughs.

> **Now we have a common language in our office of academics, our office of schools, and then all of our buildings, so the conversation has changed pretty dramatically. And we can measure instruction now. We couldn't do that before.**
>
> ***—District leader (2018)***

In half the districts, central office leaders and supervisors clearly and consistently articulated a definition of instructional leadership, at times referring to specific components of their district's principal evaluation rubric. A supervisor from Broward described a clear vision of instructional leadership and how it was tightly coupled with the district's vision of high-quality instruction:

> **[Instructional leadership means] to ensure that [the framework of high-quality instruction] happens in the classroom, to ensure that it's executed, to ensure that PLCs [professional learning communities] are held, to ensure that the data is collected and monitored so that we'll know how that data impacts what we're looking for in terms of the instructional core, what the teacher is doing, what students should be able to know and do, what curriculum is being used, what student work is being assessed and tested, and the engagement of the instruction delivery that's happening in the classroom.**
>
> ***—Supervisor (2018)***

Districts developed a more coherent approach to supporting principals and schools. The PSI increased the consistency of each district's goals for the principal supervisor role and the support that supervisors and the central office provided to schools. Supervisors and central office leaders in Broward, for example, noted improved alignment between the professional development supervisors received, the work they were doing with principals, and the language and frameworks for academic rigor used across the district. Supervisors and district leaders in Long Beach and Cleveland noted that the PSI created opportunities for a coherent, systemic approach throughout the district.

> **[The PSI] let us build true networks of support where every school now has a dedicated team that includes their principal supervisor themselves, talent partners, special education partner, family engagement partner, social emotional learning partner, academic partner. So it really let us build our robust team of people who are trained together to support schools.**
>
> ***—District leader (2018)***

One way districts helped create coherence in the principal supervisor role was by reducing the number of additional responsibilities and formal roles supervisors held in the central office, so they could focus solely on working with principals. Before and in the early years of the PSI, supervisors at times headed other departments such as research and parent engagement or were responsible for coordinating district after-school athletics. The proportion of supervisors who reported holding other formal roles in the central office fell from 49 percent in 2015 to 24 percent in 2018 (Figure IV.3). Reducing supervisors' other responsibilities allowed them to place more of a consistent focus on supporting principals.

Figure IV.3. Percentage of supervisors holding other formal roles in district declined

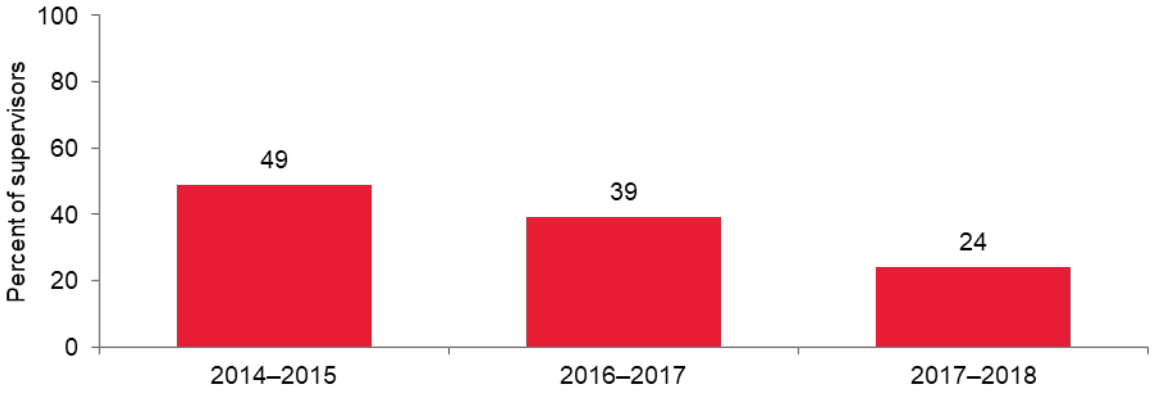


Figure reads: Forty-nine percent of principal supervisors held another formal role in the district in 2014–2015.

Source: Supervisor surveys, 2016 (*N* = 50), 2017 (*N* = 51), and 2018 (*N* = 50).

Note: The survey question read: "In addition to your role as a principal supervisor, did you have any other formal role(s) in the district during the school year (for example, director of after-school programming, liaison to the office of academics)?" The 2016 survey asked supervisors to recall their roles in 2014–2015. Change across the three years is statistically significant at the 5 percent level.

Districts modified central office structures, roles, and culture to better support principal supervisors. Over the course of the initiative, districts adjusted central office structures to facilitate communication and coordination in support of principal supervisors and their redesigned roles. These changes produced greater alignment among central office departments and helped clarify each department's specific domains of responsibility for supporting principal supervisors.

- Baltimore, for example, implemented a system of "tag teams" to facilitate communication between principal supervisors and central office departments. Each supervisor served as a liaison to at least one central office department. The goal of the tag team structure was to ensure that supervisors had input on any central office initiatives that would directly affect schools. It also created a channel for supervisors to provide feedback to the central office based on what they were seeing in schools and hearing from principals.

> **[The tag team structure] is good because it helps to build your relationship with [other departments]…when you need things quickly, or when you don't necessarily agree with something that's going on, you can go in the office, close the door and have that conversation, because you've already built that rapport with them.**
>
> ***—Principal supervisor (2018)***

- Broward held quarterly meetings and retreats to break down divisions between central office departments and align the work of principal supervisors with other departments. However, a district leader noted that although the central office had made these structural shifts, they still needed to make cultural shifts, which were far more challenging. District leaders focused on changing the mindset of central office personnel to better understand their responsibilities for supporting principals. Broward also hired a principal coach who reported to the director of leadership and was in the same department as the supervisors. Supervisors could direct the coach to work one-on-one with principals who needed additional support.

- Des Moines solidified its multiyear efforts to align and promote better working relationships between the office of schools, which oversaw principals, and the office of academics, which handled teaching and learning. The two offices were on the verge of formally merging by the end of the initiative. In addition, the district began to scale up its implementation of Schools for Rigor, a program that trained supervisors to provide more consistent feedback on student learning through a series of intensive school walk-throughs. Because Schools for Rigor was a district initiative, district leaders, supervisors, and principals in the district reported a greater degree of coherence and alignment among the central office, principal supervisors, and principals concerning the district's expectations for high-quality classroom instruction.

- Minneapolis revised its deputy chief of schools position during the course of the PSI. At first this position was focused on operational issues. In later years of the PSI, the role was revised to oversee all academic departments and coordinate their work with principal supervisors. In addition, in the first year of the PSI, the district created a turnaround support team for the supervisor overseeing the district's turnaround school network. This team was comprised of dedicated support personnel who reported directly to the turnaround principal supervisor. Turnaround personnel were assigned to work closely with the supervisor and principals of these schools to ensure they received the resources they needed. In the final year of the PSI,

the district expanded a version of this support team structure to all principal supervisors' networks by assigning central office representatives to work with each supervisor under the guidance of the deputy chief of schools. These support teams strengthened and streamlined central office support for principals and helped supervisors manage their increased spans of control.

Districts increased the rigor of their selection processes for principal supervisors. Over the course of the PSI, districts revamped their selection process for the principal supervisor role. Supervisors hired before the PSI typically reported having had one to two interviews for the position and perhaps submitting a writing sample. Supervisors hired during PSI reported a more intensive hiring process. Often they reported being interviewed by a panel of principals, some or all of the current supervisors and senior district personnel, including the superintendent. The revised selection processes also required candidates to demonstrate skills they would need as principal supervisors by participating in role-play scenarios, interpreting data reports, and providing feedback to principals after conducting a school visit with classroom walk-throughs. For example, one novice supervisor described a hiring process that included multiple interviews followed by a school visit. The visit entailed walking through the school in the same way a supervisor would and then meeting with the principal. Afterwards the supervisor candidate was required to put a plan together that demonstrated how she would prioritize her work in the school and address the principal's needs. She presented that plan to a committee that included the head of the principal supervisors' department, multiple supervisors, and the superintendent.

Aspiring principal supervisor programs became a model for broader district leadership succession planning. Broward, Cleveland, and Long Beach all developed programs to train aspiring principal supervisors in the second year of the PSI. By the end of the initiative, both Broward and Long Beach relied on these programs to fill the majority of open principal supervisor positions. Cleveland discontinued its program after learning that potential program candidates preferred to enroll in a superintendent certificate program offered by a nearby university but hired two of its graduates into principal supervisor positions. The other three PSI districts did not develop formal programs for aspiring principal supervisors.

Figure IV.4 shows the number of supervisors across all six districts who reported that they learned about the opportunity to be hired into their current position through participation in an aspiring supervisor program over the course of the PSI. By the 2017–2018 school year, one in five supervisors had been trained in an aspiring supervisor program.

Figure IV.4. The percentage of supervisors who learned about their position after participating in an aspiring principal supervisor program increased

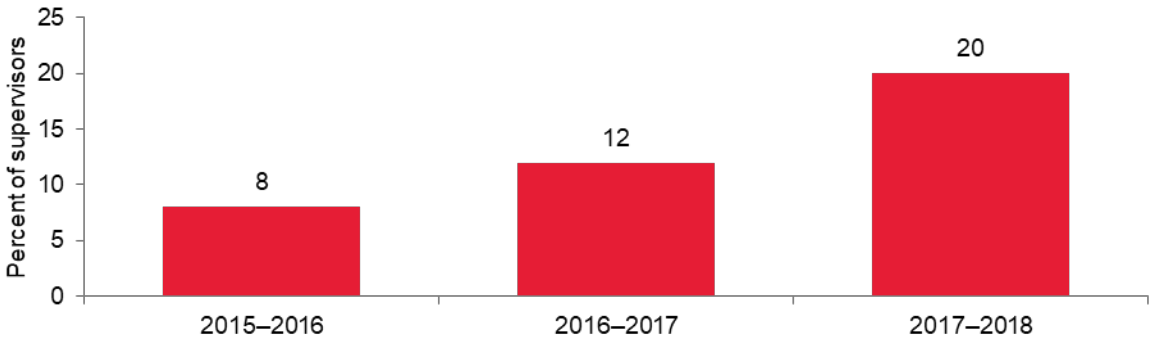


Figure reads: Eight percent of supervisors learned of their current position through participation in a program for aspiring supervisors in 2015–2016.

Source: Supervisor surveys, 2016 (*N* = 50), 2017 (*N* = 51), and 2018 (*N* = 50),

Note: Survey question read, "Thinking back to your own experience being hired as a principal supervisor in this district, did you learn of the opening for your current position through your participation in a program for aspiring principal supervisors?"

In addition, both Broward and Long Beach came to rely on the programs for succession planning for a wider variety of leadership roles in the central office. As a result, both districts broadened the scope of the programs. Long Beach implemented an additional program track for principals interested in becoming directors of central office departments, such as maintenance and facilities and student support services. Broward was planning to develop a leadership training program for future central office leaders based on the success of their program for aspiring supervisors.

There has been talk about creating leadership-type pipelines in other areas, administrative areas. Succession planning is a very challenging premise in school districts, just given a lot of the financial pressures that we face.… School systems don't readily allow you to build a number two or three in place as you know someone is going to be retiring in a couple years. In the private sector, they would spend those two years grooming a couple of folks for the position and seeing which candidate has the greatest fit and possibility for long-term success.… What we're able to do with the principal supervisor role, we're looking to create similar experiences in terms of the intern position aspect of it.

—District leader (2018)

A spotlight on the role of the superintendent in the PSI

Superintendents are key to PSI success.

The superintendent's leadership was vital to efforts to change central office and supervisor roles in the PSI districts. Superintendent leadership created a foundation for establishing and sustaining the PSI.

In those districts with the most successful implementation of the PSI, superintendents' approach to the initiative shared several common themes:

Superintendents remained engaged with the PSI work, helping to shape its direction. Superintendents varied in the degree to which they got into the "nuts and bolts" of the PSI, but in each district the leader of the PSI work reported to the superintendent, and the superintendent often weighed in on key issues related to personnel, resource use, and central office structures. Superintendents also were involved in ensuring that the PSI work connected to other district priorities.

> **[The superintendent is] very hands-on…. He stops in on a lot of things that are where the staff is working together. He meets regularly with the [principal supervisors]…. And he has come a couple times, like to our retreat setting to talk about metrics that he's looking at and hear about the work. And I keep him up to date on things we're working on.**
>
> ***—District leader (2018)***

Superintendents remained committed to most key components of the PSI. The superintendents were vocal and consistent in supporting the new role of the principal supervisor in the district. They helped secure resources to maintain lower spans of control. Superintendents were also directly involved in strengthening the central office to better support the work of supervisors and principals. For example, superintendents helped make decisions about creating network teams or realigning the work of central office units.

Superintendents played a key role in championing the PSI work and making it visible to stakeholders. One superintendent described a key function of the superintendent as elevating the supervisor role, "emphasizing regularly how important this position is in the district." The school board was a particularly important stakeholder because it controlled the budget, and sustainability of the initiative in future years depended on the school board being willing to support the work beyond the grant period. Superintendents in four districts described the importance of communicating about the PSI with the school board regularly as implementation unfolded.

> **"The other part is for me to make sure that those [school board] members understand what we're doing and so the team has presented numerous times throughout the school year on the whole initiative…. If you ask my board members, they'd be able to tell you … we know what they're doing because we've heard it here, here, and here.**
>
> ***—District superintendent (2018)***

The superintendents' vision for the PSI work following the end of the grant period appeared key to determining what aspects of the PSI would be sustained and in what form. In some districts, this meant planning and budgeting for the work to continue largely as-is after the grant period. In others, this meant focusing on specific components that the district found most valuable. One superintendent, for example, described aiming to strengthen the district's model for grouping schools into networks and the associated support structure that tied specific central office department personnel to each network. The same superintendent also emphasized a commitment to continuing to change the central office structure and culture to better meet the needs of schools.

Stability in the superintendent's office was particularly important. In the four districts with the most success in carrying out the PSI work, the same superintendent was in place throughout both the planning and implementation stages of the PSI. These superintendents understood the PSI and its goals from its inception.

Two districts experienced superintendent turnover during the PSI. Just as stability was important for implementation success in the other four districts, superintendent turnover led to change or retreat from implementation of the PSI components in these two districts. In both cases, the new superintendents' visions for the principal supervisor role did not align with the prior administrations' visions. The new superintendents advocated for changes in how principal supervisors conducted their work. In one district, the new superintendent sought to align principal supervisors' work to a different vision of instructional leadership and principal support. Whereas the previous administration had redefined the supervisor role to be purely instructional in focus, the new superintendent called for a role that balanced supervisors' instructional leadership work with operational and managerial work. In the other district, the principal supervisor role shifted more fundamentally, becoming a liaison between the superintendent's cabinet and the schools, with reduced emphasis on the principal support and coaching activities prioritized in other PSI districts. These changes led these districts away from the goals of the PSI in the final years of the initiative.

B. Principal supervisors' experiences

Principal supervisors focused on supporting principals. Over the course of the PSI, supervisors developed clarity about the purpose of their work: a combination of supporting and developing principals and monitoring school performance. By the final year of the PSI, supervisors noted they now had key strategies, frameworks, and tools to use in their work and that their job had become less haphazard and more systematic over time. They saw their work as focused on coaching and grounded in instruction, data, and asking questions.

> **Throughout the four years I have to say what has changed the most is my skillset. …[PSI] gave me more clarity about what really my role was all about, and how I could impact a more in-depth level with my principals, and how I could work differently with them.**
>
> ***—Principal supervisor (2018)***

> **There has been a stronger focus on understanding good instruction, being able to coach around it, being able to support principals and implement academic programs that reflect the rigor of the standards.**
>
> ***—District leader (2018)***

Supervisors and district leaders also reported that supervisors developed closer relationships with their principals and became much more informed about the principals and their school contexts.

> **I think there's been a huge difference in the ownership for principal supervisors of what they know about what the schools are experiencing in terms of their data, their success, the things that they are doing to really advance on student achievement in their building…I think there is an ability for our principal supervisors now to advocate.**
>
> ***—District leader (2018)***

Principal supervisors developed a shared professional identity and community of practice. Over the course of the PSI, many supervisors developed a sense of professional identity rooted in a set of shared practices and values. They felt that expectations about how much time they should spend in schools and how that time should be used were well defined. They planned together and learned from one another, often in communities of practice in which they came together to discuss problems of practice and improvement. They became the leaders and co-developers of their own professional development, whereas in the earlier years of the PSI other central office officials largely drove this.

Supervisors and central office staff in some districts felt that the opportunity to learn from other districts was another big accomplishment of the PSI. For example, Broward established an annual national conference focused on the principal supervisor role. The conference provided opportunities for supervisors from across the country to learn about approaches to the role, share how they were addressing similar hurdles, and meet other supervisors to develop professional networks that they could call upon throughout the year.

> **We started having a community of practice around instruction, where we normed our practice for evaluation, where we come together to solve similar problems versus trying to solve them individually.**
>
> ***—Principal supervisor (2018)***

Principal supervisors reported that their practices became more consistent within districts. Many districts worked to increase the consistency of practices among supervisors to ensure principals received comparable support and calibrated ratings on evaluations. Districts strived to improve consistency by sharing common goals for supervisors, standardizing the focus of school visits, adopting common tools to guide supervisors' interactions with principals, working to calibrate evaluation ratings, and setting expectations for the amount of support supervisors should provide to principals and schools. Supervisors and central office staff talked at length about how the PSI helped establish collaborative structures that enabled them to be more consistent in their practices.

> **We expect that [principal supervisors] are visiting schools biweekly, that they're in schools … 60 to 70 percent of the time … that they are visiting and doing side by side walks, classroom observations, that they are visiting PLCs, that those day to day functions of the principal … are being observed, and they're having coaching conversations and next steps so that we can develop the skills and knowledge in those particular areas. So those are common expectations for all of them.**
>
> ***—District leader (2018)***

The development and use of tools and protocols helped make principal supervisors' practices more consistent. The proportion of supervisors who usually or always used specific protocols when discussing data and making school visits increased from 2016 to 2018, although these increases were not statistically significant (Figure IV.5).

Figure IV.5. Supervisors' use of standardized tools increased

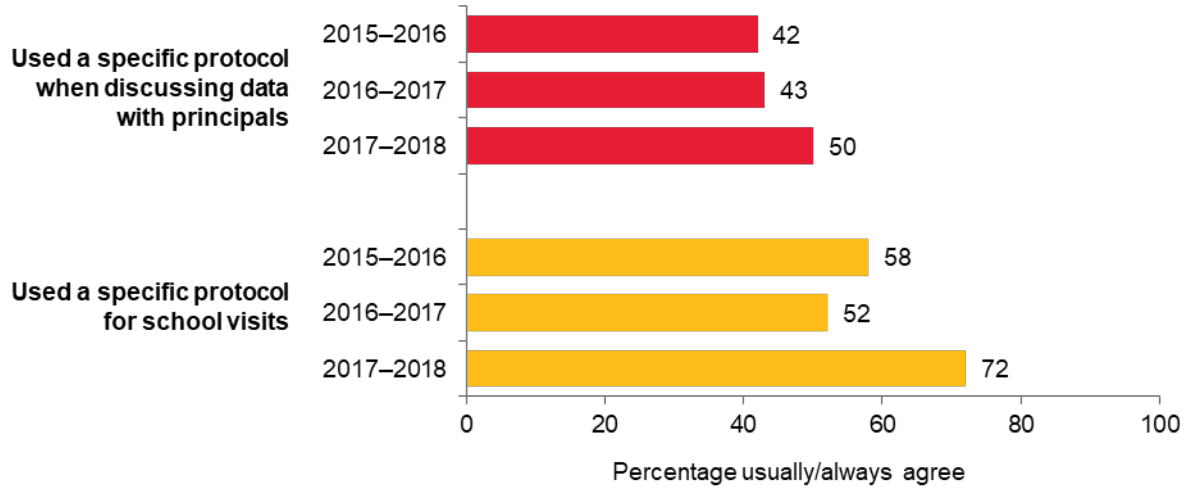


Figure reads: Forty-two percent or supervisors reported that they usually or always used a specific protocol when discussing data with principals in 2015–2016.

Source: Supervisor surveys, 2016 (*N* = 50), 2017 (*N* = 51), and 2018 (*N* = 50).

Note: The survey question read, "Over the past three months, when you visited a principal at his/her school, how often did you do each of the following?" Changes across the three years are not statistically significant at the 5 percent level for either item.

Principal supervisors spent the majority of their time with principals. In the final year of the PSI, principal supervisors spent 63 percent of their time working with principals: 48 percent of their time was spent visiting schools, and 15 percent was spent in principal group meetings (Figure IV.6). These numbers had remained high over the course of the initiative, with time spent working with principals ranging from 60 to 63 percent over the four years of the PSI. Early in the initiative, districts set expectations for supervisors' one-on-one work with principals that helped supervisors focus their days. For example, district leaders in Cleveland expected supervisors to visit each school at least once a week. Most supervisors created weekly schedules that detailed which schools they would visit and for how long, depending on the school's needs.

Supervisors reported that they became more intentional in planning and leading principal group meetings. These meetings were typically scheduled at the beginning of the year to ensure that they would be a priority. Although some cancellations were inevitable, supervisors and principals reported that principal group meetings occurred regularly over the course of the initiative. Supervisors implemented a variety of activities in these meetings with their principals, such as data discussions and book clubs focused on leadership and coaching topics. Some supervisors allowed principals to lead these network meetings and give presentations to their peers. Others rotated their meeting locations throughout the year so that each principal could host and lead the other principals on a guided walk-through of their school building.

Figure IV.6. Principal supervisors spent the majority of their time working with principals in 2017–2018

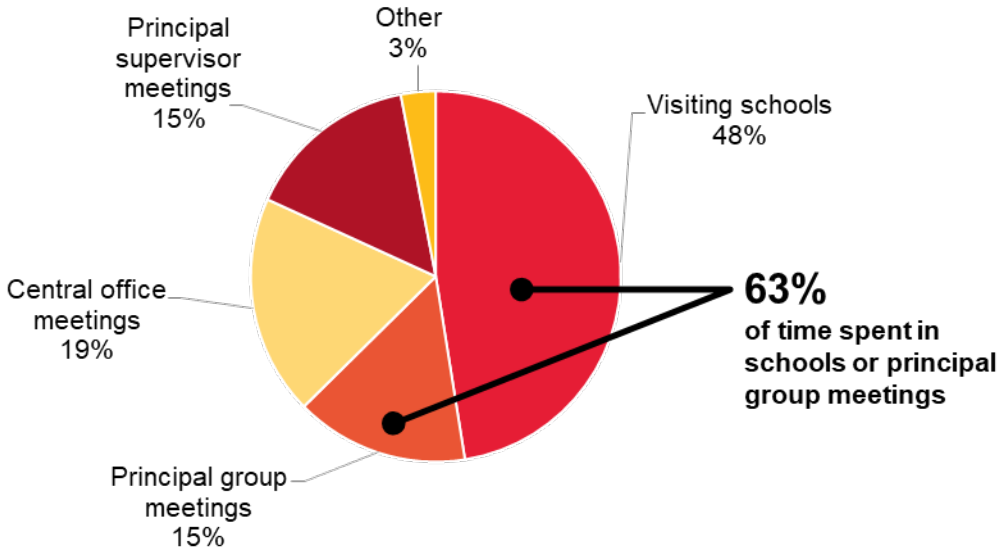


Figure reads: Principal supervisors reported spending 48 percent of their time visiting schools in 2017–2018.

Source: Supervisor surveys 2018 (*N* = 50).

Note: The survey question read, "Over the past three months, what proportion of time did you spend on each of the following activities in a typical week, excluding travel time?"

Principal supervisors received better support from the central office. Districts discovered early on that they needed to engage the entire central office to support the changes to the supervisor role. Providing coherent, unified support to principals required principal supervisors to collaborate with other central office departments and personnel. Central office personnel also needed to become more familiar with and integrated into the work of schools. Districts continued to make use of structures they had created in the early years of the initiative to integrate the central office with supervisors' work. These included cross-departmental liaison roles, where each supervisor was assigned to be a liaison to another department, such as Baltimore's tag teams. They also included district support teams, where supervisors were assigned dedicated representatives from other departments who met with them regularly to discuss ongoing work, attend their principal network meetings, and work directly with their principals such as the teams formed to support networks in Minneapolis.

Supervisors' ratings of the quality of central office support for both their work and for principals' work increased over the course of the initiative. Figure IV.7 shows a scale measure of supervisors' average agreement with 11 items that asked them about the quality of support provided by the central office, including statements such as "The district central office facilitates my work with principals" and "Improving teaching and learning is a key focus of the central office's work" (see Appendix A for information on construction of this scale). However, although they increased over time, supervisors' ratings of the quality of the central office's support suggested that it remained an area for improvement. By 2017–2018, supervisors, on average, neither agreed nor disagreed with each statement about the central office, as reflected by a score of 3.0 on the 1-to-5 scale.

Figure IV.7. Supervisor ratings of the quality of the district central office support increased

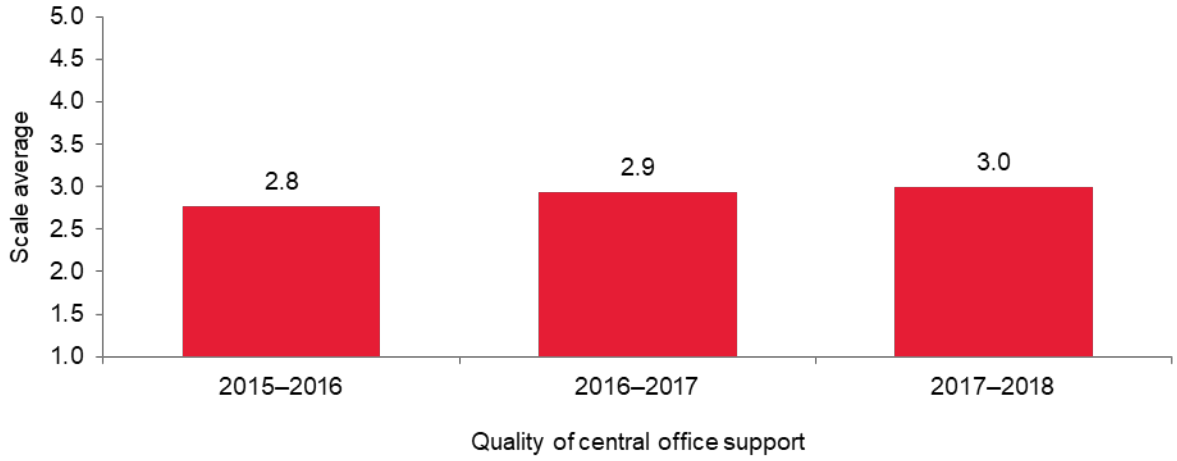


Figure reads: Supervisors' average rating of the quality of central office support in 2015–2016 was 2.8.

Source: Supervisor surveys, 2016 (*N* = 50), 2017 (*N* = 51), and 2018 (*N* = 50).

Note: Scale ranges from 1 to 5. See Appendix A for information on scale construction. Changes in ratings across the three years were statistically significant at the 5 percent level.

A Broward supervisor spoke about how the academic department became noticeably more integrated with the supervisors' work developing principals. Instructional specialists and curriculum experts actively supported schools and were involved in small group training sessions facilitated by supervisors.

> **[Before], you really weren't sure who to call and who could be supportive. I don't think we had the level of instructional specialist connected to our schools years ago, and definitely there's much more of a boots on the ground support from academics**
>
> ***—Principal supervisor (2018)***

A spotlight on supervisors' differentiation of support for principals[9]

Supervisors differentiated supports for principals

Reduced spans of control and a reorientation of the principal supervisor role enabled supervisors to individualize and differentiate supports and growth opportunities for principals based on their needs and backgrounds.

Some districts formally differentiated supports for principals. Districts often assigned their schools to specific categories according to the intensity of the principal or school needs, a process commonly referred to as *tiering support*. While some districts, like Long Beach, had a systemic approach of categorizing schools and determining patterns of differentiation, other districts had less clearly defined criteria, which gave supervisors flexibility in determining how to create and support each tier. In districts where differentiation was not strategically predetermined, differentiation of supports varied across supervisors.

Long Beach defined formal criteria and approaches for differentiating support to principals based on the principals' evaluation ratings. The district also used school-based factors, such as culture and climate or community relations, to make adjustments to the principals' assigned tier. Supervisors then provided levels of support based on the tier: Tier 1, Basic Support, at least 3–4 hours on site every six weeks; Tier 2: Essential Support, at least 4–6 hours on site every six weeks; and Tier 3: Targeted Support, at least 8–10 hours on site every six weeks (Figure IV.8). The district also developed guidelines for supervisors' on-site activities and monitoring for each tier of schools.

[9] This section was coauthored by Kristen Carroll of Vanderbilt University.

Figure IV.8. Long Beach's differentiated principal supervision guidelines

long beach unified school district

Differentiated Principal Supervision Guidelines

Pilot Year – 2017-2018

Tier I: Basic Support

The Basic Support category is associated with principals who need the least support in accelerating student achievement, may be "stars" or models for others. The Basic Support category may also include schools where expected continuous improvement has been demonstrated.

Time	Typical On Site Activities	Monitoring
3-4 hours on site each six weeks	- Admin Team and 1:1 coaching -Calibrate observations of classrooms together -Process observer for various meetings w/feedback -Could have an hour of networking w/PS as facilitator around special topics in place of time -Principal could lead a networking event (e.g. study group, book study group) to delve deeper into their own goals/action plans and plan next steps	-Quarterly review of Goals/Action Plans -Quarterly review of Interim Data -Bi-annual review of teacher supervision and evaluation documents -Regular classroom visits -Regular review of site budget and spending

Tier II: Essential Support

The Essential Support category is associated with principals who need some strategic support in ensuring the implementation of their school goals/action plans or have made leadership progress, but have not reached the basic support level. The Essential Support category may also include schools where continuous improvement has not met district expectations in pace or outcomes.

Time	Typical On Site Activities	Monitoring
4-6 hours on site each six weeks	-Admin Team and 1:1 coaching -Calibrate observations of classrooms together -Calibrate observations of classrooms with Curriculum staff -Process observer for various meetings w/feedback -Review teacher evaluation process documents -Strategic focus on Teaching & Learning along with Strategy & Planning -Assist in monitoring	-Monthly review of Goals/Action Plans -Monthly review of Interim Data -Monthly review of teacher supervision and evaluation -Frequent classroom visits -Frequent review of site budget and spending

Tier III: Targeted Support

The Targeted Support category is associated with principals who need the highest level of support in implementing their school goals/action plans and who are considered novice, struggling, inconsistent or low performing. The Targeted Support category may also include schools where there is a lack of continuous improvement and/or limited progress.

Time	Typical On Site Activities	Monitoring
8-10 hours on site each six weeks	- Admin Team and 1:1 coaching -Visits may include additional central office staff -Calibrate observations of classrooms together -Calibrate observations of classrooms with Curriculum staff -Process observer for various meetings w/feedback -Review teacher evaluation process documents -Strategic focus on Teaching & Learning along with Strategy & Planning -Assist in monitoring -Development, review or monitoring of ISP (as applicable)	-Monthly review of Goals/Action Plans -Monthly review of Interim Data -Monthly review of teacher supervision and evaluation -Frequent classroom visits -Frequent review of site budget and spending -Monthly review of ISP (as applicable) -Principal will be assigned an additional coach to support performance improvements

In other districts, supervisors used more individual discretion to assign principals and schools to tiers. Student achievement, principal evaluation scores, and principal years of experience were all common criteria supervisors reported using to place principals into support tiers. For example, Cleveland set expectations about criteria for tiering, which included student achievement, teacher evaluations, and principal evaluation scores. Although Broward and Des Moines did not have districtwide criteria for tiering, supervisors were expected to tier schools based on their own discretion. In Des Moines, leadership deliberately left the tiering process to supervisors with the expectation that differentiation should be based on a "thorough analysis and assessment based on evidence" of which principals are meeting expectations. In Broward, supervisors operated under the clear expectation that they should tier their schools, but specific strategies were left to the individual supervisors.

> **I can tier those schools, so I can determine where I'm going and how much time I'm going to spend at the most needed schools…. It's based on a lot of data, multiple sources of data, attendance rates, teacher success and their impact on classes… whether principals have shown or not shown improvement in certain, particular areas that I think contribute to the success of a school.**
>
> ***—Principal Supervisor (2018)***

Supervisors differentiated time, allocation of resources, and task emphases

Supervisors spent more time with less experienced and lower-performing principals. The most common form of differentiation was in the allocation of time and the number of visits supervisors made to schools. The majority of supervisors contacted their principals once every two weeks. However, supervisors reported that they spent more time in schools working with principals when there was greater need, such as lower-performing schools or schools "in crisis."

> **One of my schools is in absolute crisis mode right now. I'm there at least two to three times a week. I have a principal retiring at the end of this year. So I've been … transitioning a new principal in. I've prioritized that school. I've also released a few schools. They've been showing good progress, good updates. They have data, and the principal and I can talk on the phone and she'll send me a few things. And so I'll hit them once maybe every week and a half or two weeks. Just the releasing of the depth of support needed—if they just need a check-in to see. The ones I'm prioritizing need me.**
>
> ***—Principal Supervisor (2018)***

Similarly, supervisors visited less experienced principals more frequently (see Figure IV.9). Supervisors reported visiting with experienced principals an average of six times in a three-month period. However, supervisors visited less experienced principals an average of seven times over a three-month period.

Figure IV.9. Supervisors visited their less experienced principals more frequently

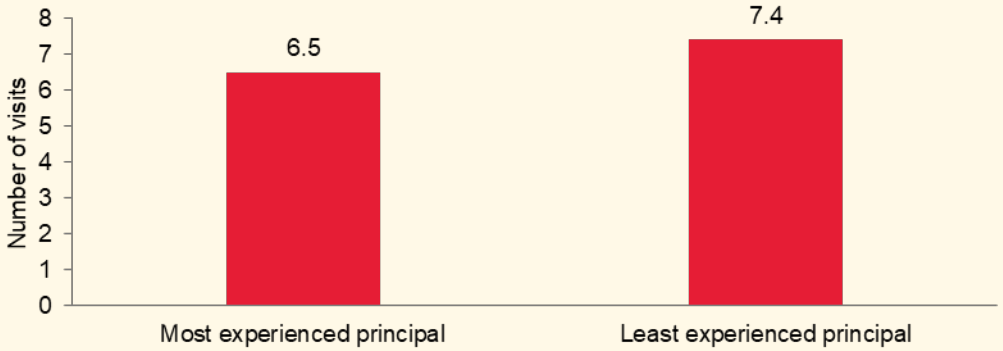


Figure reads: Supervisors visited their most experienced principal's school 6.5 times on average and visited their least experienced principal's school 7.4 times on average over a three-month period in 2017–2018.

Source: Supervisor surveys, 2018 (*N* = 50).

Note: The survey question read: "How many times have you visited [school] in the last three months?" Differences between the two groups are statistically significant at the 5 percent level.

Supervisors placed greater emphasis on instructional leadership work more with their least experienced principals. Principal supervisors differentiated their practices in response to principals' years of experience. Supervisors reported in interviews that they tended to spend more time with their least experienced principals, providing both instructional and noninstructional support. In surveys, we asked supervisors to report their emphasis on key practices with their most and least experienced principals. Supervisors reported that on average they placed more emphasis on developing principals' instructional practices (for example, guiding the principal in developing schoolwide professional development), feedback practices (for example, modeling or role-playing teacher feedback), classroom visit practices (for example, working with the principal to give feedback based on classroom visits), and data use practices (for example, helping the principal think through the school's data) with the least experienced principal in their network than with the most experienced principal. Figure IV.10 shows that 70 percent of supervisors placed more emphasis on developing principal's feedback practices with their least experienced principal compared with their most experienced. By comparison, only 23 percent of supervisors placed more emphasis on feedback with their most experienced principal. Similarly, 57 percent of supervisors placed greater emphasis on instruction practices with their least experienced principal.

Figure IV.10. Supervisors emphasized instructional leadership practices more with less experienced principals

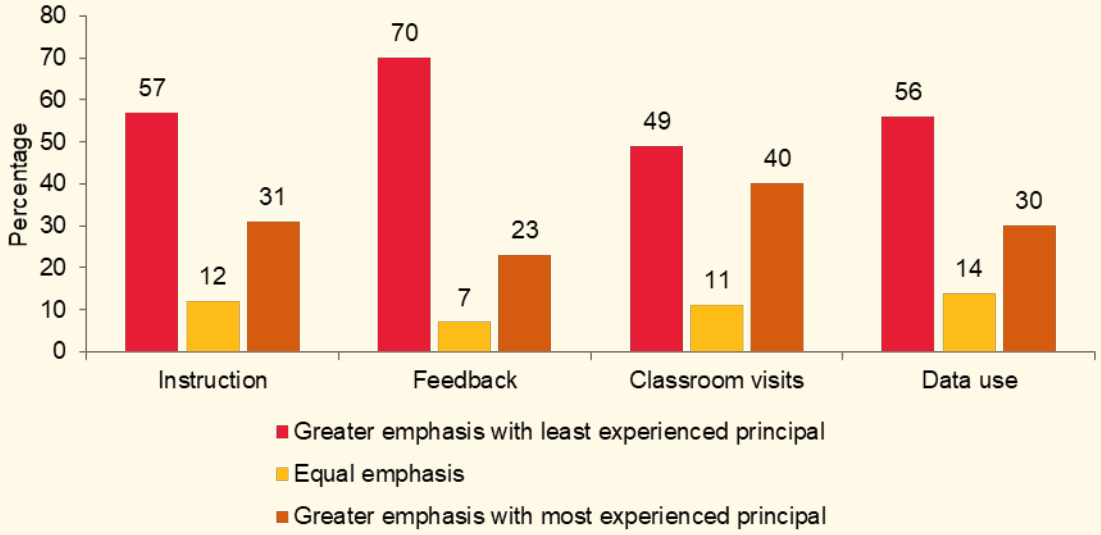


Figure reads: Fifty-seven percent of supervisors reported placing more emphasis on instruction practices with their least experienced principals in 2017–2018.

Source: Supervisor surveys, 2018 (*N* = 50).

Note: Multiple survey questions were included in each practice category (see Appendix A for information on scale construction).

We might be able to walk two or three classrooms with a veteran principal and I get a sense that they understand instruction, they understand the feedback, they have a really good sense of how their professional development is connecting to their observation and feedback and others might need a little bit more time to do that. So it could be a differentiation in the amount of time that we're spending there. And then I think just ultimately because everything we do is grounded in their evaluation system and their evaluation rubric… So it's their needs based on their experience and where they are in their own proficiency as a principal.

—Principal Supervisor (2018)

Furthermore, supervisors often spent time helping new principals with school climate issues, the school budget, and developing school improvement plans.

Supervisors, particularly in Cleveland and Long Beach, reported that the style and focus of their coaching was individualized according to principals' needs. They provided more directive coaching to lower-performing principals, suggesting what needed to be done and how to do it. Their coaching of higher-performing principals was more facilitative, often resembling a conversation where the supervisor served as a "thought partner."

Supervisors directed supports, such as instructional coaches, to the schools with the greatest needs. Another mechanism to differentiate supports was via strategic deployment of support personnel such as instructional coaches to schools. In four districts, Baltimore, Broward, Cleveland, and Des Moines, supervisors used this approach to direct supports to the schools with the greatest needs. Similarly, Long Beach also allocated coaches to the schools deemed to have the greatest needs, but decisions about how coaches were allocated were made by district leaders under the direction of the superintendent.

Differentiation did not address the ongoing development of high-performing principals. Principal supervisors differentiated their attention among their principals based on school contexts. Schools with low achievement or those with novice principals received the most support and attention. Principals perceived as high-performing, who led high-achieving schools, or principals that had few "fires" to put out, received much less support and attention. These principals typically interpreted this as a sign that they were satisfactorily meeting expectations. For most districts, this approach to differentiation did not result in ongoing development to challenge, grow, and engage more experienced and higher-performing principals in their own leadership development or in mentoring them toward district-level leadership.

C. Challenges faced by districts and supervisors in changing the principal supervisor role

At the end of the initiative, districts reflected on the challenges they faced in changing the supervisor role over the four years of the PSI. These challenges included reducing supervisors' span of control, fostering structural and cultural changes in the central office, and balancing all the responsibilities placed upon supervisors while also finding time for ongoing training and collaboration.

Reducing the span of control for all supervisors was difficult. Reducing the span of control required either new resources or reallocation of current resources to hire additional principal supervisors. For some districts, garnering these resources was extremely difficult, and not all supervisors experienced a reduction in the number of principals in their networks. Although the average span of control across all districts in the final year was 13 principals, spans ranged from 7 to 19 principals (Figure IV.11). The PSI encouraged districts to ensure that the positions remained sustainable, with adequate funding, after the PSI funds were no longer available. This was a more expensive endeavor for districts, which needed to hire many supervisors to bring down their span. Some district leaders noted that adding costly positions to the central office could be politically contentious.

Figure IV.11. Span of control varied in the PSI's final year

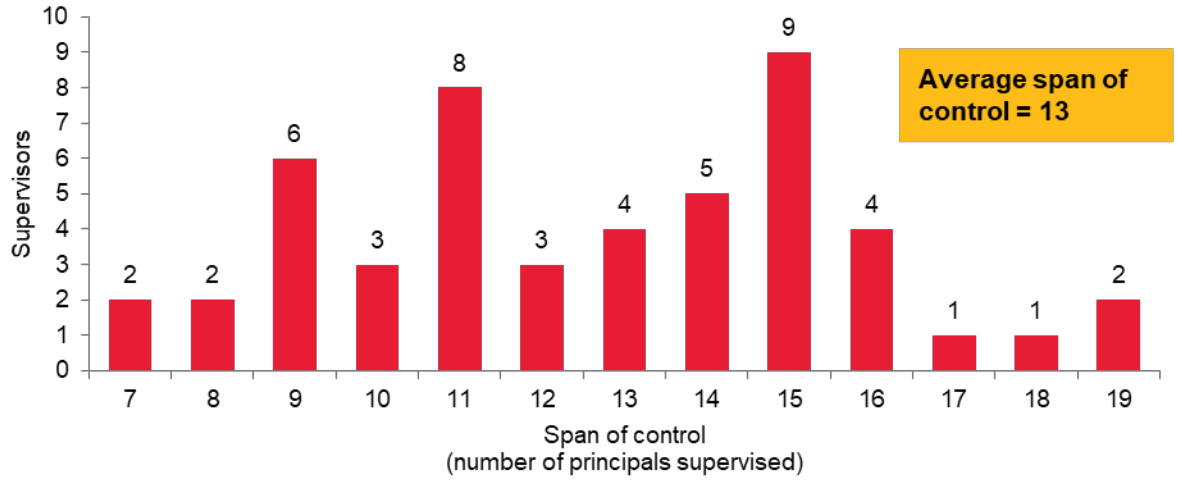


Figure reads: In 2018, two supervisors oversaw 7 principals, two oversaw 8 principals, six oversaw 9 principals, and so on.

Source: Principal supervisor survey 2018 (*N* = 50).

Note: The survey question read: "How many principals do you currently supervise?"

The span of control varied for individual supervisors because each district differed in how it assigned principals to supervisors (for example, by geography, grade level, school performance, or school theme such as arts or science, technology, engineering, and mathematics [STEM]). In one district, for example, turnaround school networks were capped at 11 schools. The other schools in this district were grouped by grade level, resulting in a span of 19 schools for one supervisor. In another district, networks based on themes varied in size according to the number of schools with each theme, resulting in spans ranging from 10 to 14 principals per supervisor.

Finally, turnover and vacancies in the supervisor position forced districts to reassign schools to other supervisors, creating higher than usual spans of control for some supervisors.

Aligning central office culture and structures with the new principal supervisor role was challenging. The new principal supervisor role required supervisors to work closely with other central office departments. Staff from most districts said it was a heavy lift to reframe the work and change the culture of the central office to better support schools and principal supervisors. District staff and supervisors found that it was difficult to obtain buy-in from central office personnel and to deepen communication, consistency, and coordination across multiple central office departments with different goals and operating styles. It was also difficult to determine how to reallocate tasks no longer carried out by supervisors.

> **I would say the heaviest lift was in getting long-time central office people to think differently about what it means to be on a team that supports schools. … We spent an inordinate amount of time in getting to where we are today, which isn't near far enough, and getting people to think of their role as service to schools instead of processing the payroll.**
>
> ***—District leader (2018)***

Some central office departments were resistant to change. The PSI districts devoted tremendous effort to overcoming resistance to change in the central office, but some departments remained more resistant than others. Some central office departments were unresponsive, made demands of principals that failed to take into account their many other responsibilities, and complained when their requests were not treated as principals' top priorities. Across districts, principal supervisors reported that helping principals navigate unhelpful, unresponsive central office staff detracted from their work with principals on instructional leadership.

The new expectations for the supervisor role were demanding. The PSI required many supervisors to make substantial shifts in their daily work. District leaders noted that some supervisors were resistant to change. Some supervisors preferred the compliance-oriented portion of their role that the PSI reduced. As the role shifted, these supervisors found themselves unsatisfied by the new demands and expectations for a role they had grown accustomed to. Some supervisors chose to leave the position rather than endure the shift.

> **If you are just used to being compliant, you had to build skills and a muscle that you don't normally use … they had to build the instruction knowledge to be able to do this work. So that was the heavy lift.**
>
> ***—District leader (2018)***

It was difficult to sustain high-quality supervisor training and professional development over the longer term. Over time, supervisors noted that dedicated training opportunities became less valuable or stale. Some supervisors attributed this to a shift away from external technical assistance providers. Newer supervisors did not receive the same strong, foundational training as their colleagues had received in earlier years of the initiative, and sometimes they did not receive onboarding at all because the training was no longer available for supervisors hired in the later years of the initiative.

On surveys, supervisors rated the quality of their training in 19 areas, such as whether the training helped build a learning community among supervisors and addressed real challenges supervisors faced in the role. They also rated the amount of emphasis their training placed on 6 concepts that were aligned with PSI goals, such as skills for coaching principals and determining protocols and procedures for school walk-throughs. When combined into scales, the ratings of supervisor training quality decreased slightly over the course of the initiative, as shown in Figure IV.12. The small decline in the scale measuring alignment of the training with PSI goals was not statistically significant.

Figure IV.12. Supervisor ratings of training quality and alignment with PSI goals declined

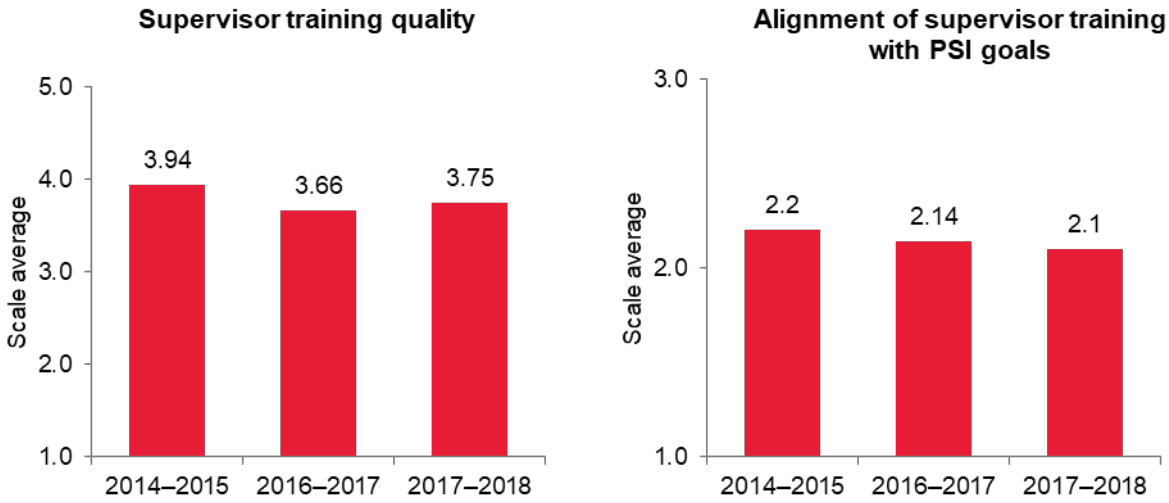


Figure reads: Supervisors rated their training quality an average of 3.94 on a 5-point scale of agreement. Supervisors rated the alignment of their training with PSI goals an average of 2.20 on a 3-point scale of emphasis.

Source: Supervisor surveys, 2016 (*N* = 50), 2017 (*N* = 51), and 2018 (*N* = 50).

Note: The training questions on the 2016 survey asked supervisors to recall training experiences in the previous 2014–2015 school year. Scale ranges from 1 to 5. See Appendix A for information on scale construction. Changes in ratings across the three years were statistically significant at the 5 percent level for ratings of supervisor training quality but not for alignment with PSI goals.

Drop-off in the quality or frequency of formal training pushed supervisors in some districts to look for new opportunities to grow professionally. Some supervisors in one district reported that their formal learning time had stagnated in the final year of the PSI. These supervisors felt that meetings had become too focused on delivering information and not focused enough on looking at data, problem solving, and sharing protocols and practices. Still, supervisors

demonstrated that they valued opportunities for training and collaboration when they deliberately set out to redirect and refocus their own learning near the end of the final year of the PSI.

> **[Other central office department] people were perceiving, 'Oh, we can get all the supervisors together in one room' [and] we let people come in and take over our agenda, and pretty soon the PLC becomes a team meeting….**
>
> ***—Principal supervisor (2018)***

Similarly, when training in one district ceased entirely, one supervisor noted that the supervisor team got "lost with [our own] development." They strove to share and collaborate informally, sometimes catching one another in the hallway to ask for feedback.

Summary. The six districts in the PSI made many changes to the central office structures and culture and to the work of principal supervisors. The districts made specific changes to the supervisor job description, such as increasing expectations for the amount of time supervisors should spend working directly with principals. Districts reduced the average span of control for supervisors by hiring additional supervisors to reduce the size of each principal network. Districts also changed central office structures and roles to facilitate principal supervisors' work. Together, these changes helped districts create more coherent visions and systems for school support and helped supervisors provide more frequent and tailored support to principals. Supervisors developed closer relationships with their principals. Greater coherence and consistency of their work also fostered supervisors' sense of professionalism and collective identity. These district changes were challenging to implement and sustain within each district's unique context but proved vital to ensuring that districts were able to successfully meet the PSI's goal of changing the role of principal supervisors.

THE FINAL YEAR OF THE PSI: MOVING TOWARD AND AWAY FROM THE GOALS OF THE INITIATIVE

The six participating districts continued to implement and refine the components of the PSI in the final year of the initiative. Although all districts made progress in different areas of the PSI, two of the districts faced weighty challenges in the PSI's final year that acutely impeded their progress. In all districts, local needs and broader district context influenced the PSI's trajectory. In cases where the districts continued their progress toward the goals of the PSI, they did so by tightly aligning the PSI's goals with other district priorities and initiatives. The PSI became integrated into the ongoing work of the district, and the needs of the district shaped the focus of the PSI work.

Districts made continued progress toward the goals of the PSI

All six districts further developed or refined their implementation of the PSI in the initiative's final year and worked toward sustaining important aspects of the initiative. Districts focused on different components and goals.

Some districts continued to make structural changes to the central office to better support principal supervisors. Districts continued to look for ways to strengthen the quality of central office support for principal supervisors' work.

- Minneapolis merged its offices of schools and academics into one Office of Academics, Leadership, and Learning. District leaders believed that this change would break down historic silos and facilitate coordination of principal supervisors' work with the 13 school support departments (for example, special education, and teaching and learning). The district also revised the role of its deputy chief of schools (the title becoming the deputy chief of academics, leadership, and learning) to focus on coordinating resources and services across the central office to support principal supervisors' work. The district also created central office support teams for all supervisors.

- In Broward the director of leadership development, who worked with principals and other leaders throughout the district, changed departments and joined the Office of School Accountability and Performance, which was also home to the principal supervisors. This helped to align the work of supervisors with other leadership work in the district, including that directly related to principal supervision. For example, the role of the director of leadership development necessitates regular interaction with the principal supervisors mainly through the aspiring principal supervisor program and through the various principal pipeline and coaching programs that the director runs. Directors collaborate with the principal coaches around specific principals' supports as needed.

Some districts aligned the work of principal supervisors with key district initiatives

- Supervisors in Des Moines were closely involved in the district's move toward the Schools for Rigor instructional model. This model encouraged principals to take a specific approach to walk-throughs, coaching, and data monitoring to better support academically demanding teaching and learning. Schools for Rigor had been implemented in a small number of demonstration schools in the 2016–2017 school year. In the PSI's final year, principal supervisors were instrumental in expanding the model throughout the district.

- In Long Beach, principal supervisors were closely involved in the district's efforts to improve equity, including closing the achievement gap for African American students and better meeting the needs of English language learners. Principal supervisors formed Quality Improvement Teams to study these issues. They then developed, tested, and implemented new approaches for their work with principals, including a series of questions to use during walk-throughs that specifically addressed issues related to student equity.

- In Cleveland, principal supervisors trained principals on the district model and rubric for high-quality instruction. Supervisors began leading principal roundtable meetings, which were formerly led by the chief academic officer. The supervisor-led meetings began in summer 2017 and continued through the 2018–2019 school year (a year after the PSI ended). They included all principals and assistant principals in the district. In the meetings, supervisors, principals, and assistant principals reviewed and critiqued student work and discussed implications for quality of instruction. Supervisors in Cleveland also worked with principals on providing clear and actionable feedback to teachers and sought to incorporate this same type of feedback into their own coaching with principals.

- In Broward, the district made changes to principal training to better use principal supervisors' expertise. In the early years of the PSI, supervisors brought all their principals together for group meetings, known as cadre meetings. They then grouped principals within their cadres into subcadres based on specific needs or similar characteristics (for example, a focus on STEM) for additional regular meetings. In the final year of the PSI, the district redesigned its subcadres to include assistant principals, literacy coaches, and two teachers from each school to provide more coherence and depth to the training.

- In Baltimore, the district continued the tag team system in which each supervisor served as a liaison to at least one central office department. This system provided a channel for supervisors to share principals' and teachers' perspectives with central office staff. Supervisors were also better able to obtain integrated information and viewpoints from the central office to share with each other.

Districts aligned supervisor training and other support opportunities with districts' priorities and needs. In the first three years of the PSI, all six districts relied on external technical assistance providers to develop and facilitate professional development and support for supervisors. However, most districts decreased or eliminated their reliance on technical assistance providers by the end of the initiative. Often districts continued using the structures and practices introduced by technical assistance providers (or co-created with the district) even after the providers left the districts. These structures and practices included coaching models, protocols for classroom walk-throughs, and lab days that brought groups of supervisors (and at times other district-level personnel) together to examine supervisor practice in action in schools. Furthermore, the training in some districts shifted in the PSI's final year to focus more intensively on the supervisor's work and to respond to changing district priorities that required supervisors to develop new skills, such as deepening their understanding of updated state learning standards. Some districts worked to identify the individuals who would be responsible for facilitating supervisor training after the end of the grant.

- Training in Long Beach allowed supervisors to become the leaders and co-developers of their own professional development, whereas in years past the deputy superintendent of schools largely drove this training. The general focus moved beyond how to do the principal supervisor job to focus on sharing and refining supervisors' practice. The training placed a heavy emphasis on consistency of practices across supervisors, especially in evaluation practices.

> **A number of [supervisors] participate in a planning group … and are really a part of the agenda development for the meetings… Because we have more systems in place now, that meeting is much more organic and is about sharing practice and not just about the logistics of doing the work.**
>
> ***—District leader (2018)***

- In Des Moines, supervisors went on "Rigor Walks" where they were trained to diagnose aspects of classroom instruction that needed improvement and to coach principals based on their observations.

> **I had to be calibrated with the trainers…walk a building and make sure our scoring was spot on… that kind of helps me with the coaching piece, knowing that you have to know what good instruction looks like… that's been the big part in the investment into my training, so then I can help coach principals, assess their building, and help them with next steps.**
>
> ***—Principal supervisor (2018)***

- Baltimore laid the foundation to expand its principal supervisor trainings in the future. It established Friday Collective Learning, a structured time for training where supervisors and the chief of schools allotted designated time to work on developing skills.

- In Cleveland, supervisors shifted the content of their own training toward a greater focus on ensuring high-quality instruction in schools. The supervisors focused on writing instruction, in particular, after the district identified it as an area of need based on previous student test scores.

Some districts moved away from some of the goals of the PSI

In the final year of the initiative some districts moved away from some of the goals of the PSI as they faced internal challenges that led to a drop-off in PSI implementation. In one district, a budget shortfall reduced resources and led to a downsizing of central office personnel, including principal supervisors. The district also hired a new superintendent whose vision for the principal supervisor role differed from the vision set forth by the previous administration during the district's initial implementation of the PSI.

Another district made changes to its central office leadership and structures. Supervisors were moved from the academics office and into a new office with a new chief. This change required a shift in how supervisors and the new chief coordinated their work with the rest of the central office. The two departments worked to determine which had the expertise and oversight to make suggestions, recommendations, and changes for principals and schools.

Two districts found it difficult to maintain a low span of control for supervisors. In one district, supervisor spans fluctuated throughout the year due to supervisor turnover and position vacancies. Another district eliminated two supervisor positions due to budget shortfalls. Furthermore, toward the end of the 2017–2018 academic year, this district eliminated all four principal supervisor positions. The supervisors were told they could reapply for one of only three supervisor positions in the following year. The reduction in positions led to reassignment of principals to new, larger supervisor networks. In this district, supervisors expressed reservations about reverting to this "older model," fearing that the high span of control would make it difficult to adequately support schools.

The increased span of control for some supervisors in these two districts contributed to principals' perceptions that their supervisor's span of control interfered with their ability to provide them with adequate support. In these districts, principals were most likely to report that their principal supervisor supervised too many principals to provide them with enough support. On average across all six districts, 13 percent of principals agreed or strongly agreed with this statement, compared with 34 percent and 19 percent in the two districts with higher spans of control.

In two districts, supervisor turnover and assignment changes created instability. Supervisor turnover in two districts spiked in the final year of the PSI, leading to a high percentage of principals with new supervisors.

Supervisors and principals across all six districts noted that stability of principal supervisors and the schools in their networks allowed supervisors to get to know the schools and principals and to build continuity, trust, and rapport.

> **We just don't do consistency well as a system…. There's such power in consistency, and I realize it even more with having two brand new schools this year, and just that process of … getting to know the bare basics of this school, and to think about the level of knowledge that I have about my other schools and what layer of questions I'm asking. Because schools are like onions. You just keep peeling away and understanding them more and more and more until you kind of get to the root of like, a-ha! Here's what we got to tackle right now. The consistency is everything.**
>
> ***—Principal supervisor (2018)***

In two districts, some supervisors reported a shift in focus from instructional leadership back to operations. Vacancies and downsizing in the central offices in two districts reduced supports for principals and supervisors. Supervisors in these districts consistently reported that vacancies in the central office required them to shift their focus away from instructional leadership. Instead, they spent significant time on operational tasks. Some support personnel positions that were previously assigned to assist principal supervisors in their networks were eliminated or went unfilled, which in turn shifted noninstructional work back onto principal supervisors.

> **I believe at this point, there are maybe three, maybe four building managers for the entire district. So we share. I think my building manager has four networks, has about 50 schools or something like that…. People left … and they haven't been replaced…. The manager kind of did everything else related to the school outside of academics. So the facilities issues, all of that, whatever needed to be done.**
>
> ***—Principal supervisor (2018)***

In one district, supervisors described a shift in their focus away from coaching principals. They were expected instead to handle everything related to principal and school support, such as the operational issues that had been removed from their responsibilities in the earlier years of the PSI.

Many districts found it difficult to maintain programs for aspiring supervisors when there were few supervisor vacancies. The PSI districts had relatively few principal supervisor positions in the PSI's final year and difficulty predicting openings. Training people for a role that they might never have the opportunity to fill was a concern across all the districts.

Of the three districts that had developed programs for aspiring supervisors, one eliminated its program in the final year of the PSI. For two years, one district had an internally developed program for aspiring supervisors, but in the final year it disbanded the program after realizing that aspiring supervisor candidates within the district preferred to enroll in a superintendent licensure program at a nearby university.

At the same time, some district staff expressed a reluctance to take the best principals out of schools to staff the supervisor position, when effective principals were similarly in high demand. Similarly, districts also grappled with the need to attract and retain effective supervisors in a job that was often viewed as a stepping stone to other leadership positions, such as a superintendency in another district.

V. PRINCIPALS' EXPERIENCES WITH THE PSI

The ultimate goal of the PSI was to improve principals' instructional leadership by changing the culture and structure of the central office and the role of the principal supervisor. Changes to the central office were intended to support the changes to the principal supervisor role, enabling supervisors to work with principals to help implement high-quality instruction and instructional leadership in schools.

Overall, principals reported positive changes as a result of the PSI. According to interviews with central office leaders, principal supervisors, and principals across all six districts in 2015, supervisors rarely visited schools before the PSI. When they did, they typically interacted with principals solely for the purpose of oversight and evaluation. By the end of the PSI, principals across the districts described numerous changes in their interactions with and perceptions of the central office and their supervisors. They also described how these changes influenced their own practice. Principals reported improvements in supervisors' practices and skills. Principals appreciated knowing they could go to their supervisors when they needed help and in many instances said that the departments within the central office had become more supportive of schools. Across all six districts, principals discussed shifts in their supervisors' practice. They also reported that their supervisors' emphasis on high-quality instruction and instructional leadership skills was influencing their own work and helping them to become stronger instructional leaders in their schools. The changes districts made to the central office and the principal supervisor role over the course of the PSI influenced principals' views of their own leadership.

> **In the past [the district] just wanted to tell you everything you did wrong, not to share with you how you could maybe do that better, but just like kind of beat you down. And this transition into where we are now as a district, for me, has showed me that there are people out there who really want us to grow.**
>
> ***—Principal (2018)***

A. Principals' experiences with the central office during the PSI

Principals' perceptions of central office support improved during the PSI. Districts invested considerable effort to change central office structures and culture to ensure departments focused on supporting schools and principals during the PSI. Principals' perceptions of the central office improved over the four years of the initiative, although there was still room for improvement. From 2016 to 2018, there was an increase in the percentage of principals who agreed that the central office was efficient at providing services, proactive in helping meet their needs, and organized to support them (Figure V.1).

Figure V.1. Principals' perceptions of central office support improved

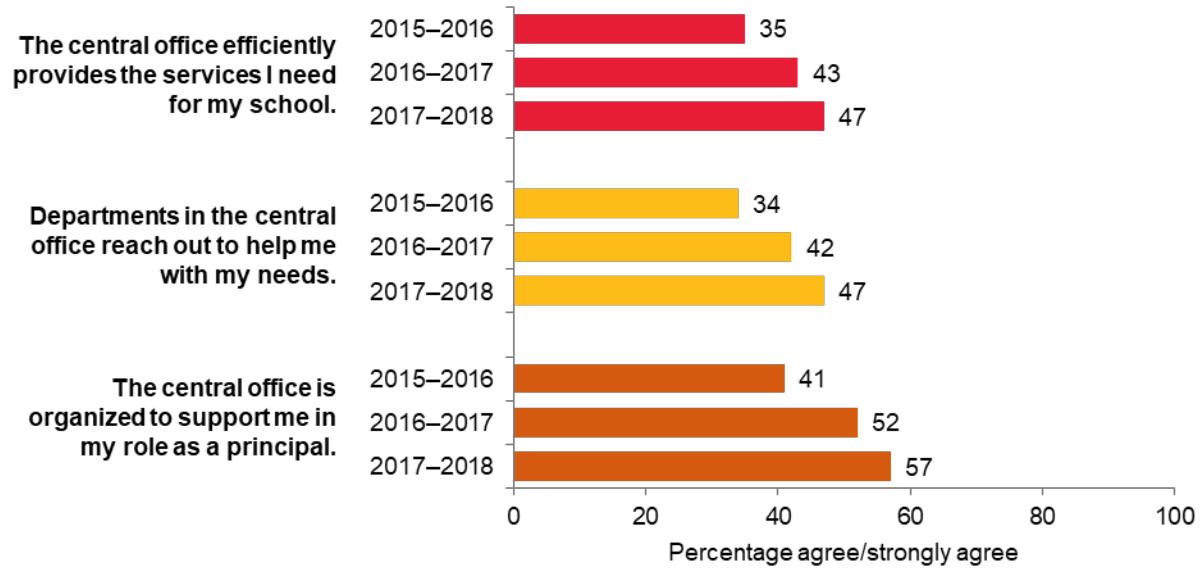


Figure reads: Thirty-five percent of principals agreed or strongly agreed that departments in the central office efficiently provide the services they need for their schools in 2015–2016.

Source: Principal surveys, 2016 (N = 635), 2017 (N = 639), and 2018 (N = 606).

Note: The survey prompt read, "Please indicate to what extent you agree or disagree with the following statements." Changes across the three years are statistically significant at the 5 percent level for all three items.

Principals used words like *consistent* and *reliable* when speaking of the positive changes that were taking place in the central office. Some principals also noted that supervisors and the central office had grown more coordinated in their approach to school support over time. Principals relied on new structures districts put in place during the PSI to get support from the central office. Staff from different central office departments were placed on support teams that assisted specific networks of schools. This structure allowed principals to reach out directly to department representatives for support (for example, a principal could contact a designated member from the budget department rather than send a message to the department as a whole).

Principals from multiple districts mentioned the increased coordination between the principal supervisors and the departments overseeing curriculum and instruction, noting increased alignment between academic initiatives and professional development for principals and teachers. Some districts also streamlined communication to ensure principals were receiving consistent messaging from various departments, which principals said they appreciated. Increased visibility of other central office departments contributed to principals' sense that the central office was becoming more accessible and was oriented toward supporting their work.

> **I think our superintendent has been really successful in getting departments to focus on schools…. I think that the focus has been more on what's going on at schools rather than what's going on [at the central office].**
>
> ***—Principal (2018)***

> **Support has become more structured and less structured at the same time, if that makes sense. And I like that. I feel like I can do my job, I have the support.**
>
> ***—Principal (2018)***

The PSI helped clarify for principals what supports were available to them from their supervisor and the central office. Access to more structured, visible support and resources helped principals become more self-sufficient in their role. Some principals reported that the increased structure in the central office actually provided them with greater autonomy because avenues for support were much clearer than in the past.

Although principals' perception of the central office improved markedly over the four years of the initiative, large numbers of principals still disagreed that the central office was efficient, proactive, or organized in its support for principals. Despite the PSI districts' progress toward the goal of changing the central office to better meet the needs of principals and schools, principals' survey responses showed that this work was far from complete.

B. Principals' experiences with supervisors during the PSI

Principals spent time working with their supervisors on instructional leadership. Principals reported that their interactions with their supervisors were largely focused on instructional leadership. They said that they valued these one-on-one interactions with their supervisors because they usually addressed real problems they faced on the job. For example, principals in Cleveland described ongoing interactions with their supervisors in which they role-played coaching conversations that they would later have with specific teachers in their schools. Principals in Long Beach noted they were expected to "know teaching and learning inside [and] out." Their work with their supervisors ensured they had the supports, training, and clear expectations about how to accomplish that aim. One principal from Long Beach described how he and his supervisor observed a classroom together, each independently wrote feedback for the observed teachers, and then sat together and compared their feedback. He called this experience "super powerful."

Over a three-month period, principals reported spending an average of three and a half hours (210 minutes, or about 54 percent of their total time spent with their supervisor in their school) with their supervisors focusing solely on instructional leadership (Figure V.2). However, there was variation in time spent on instructional leadership across districts. In the final year of the initiative in Des Moines, principals spent more than five hours (339 minutes, or about 63 percent of their total time with their supervisor) in three months focusing on instructional leadership with their supervisor. In Minneapolis, the average was slightly more than two hours (131 minutes, or about 35 percent of their total time with their supervisor).

Figure V.2. Principals' time spent working with supervisor on instructional leadership varied by district

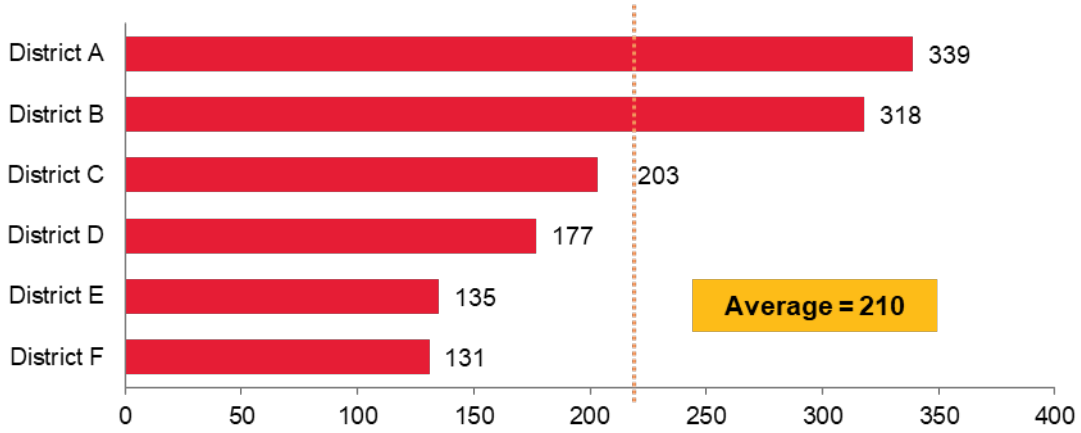


Figure reads: Principals and principal supervisors in District A spent an average of 339 minutes over three months working on instructional leadership in 2017–2018.

Source: Principal surveys 2018 (*N* = 606).

Note: Time on instructional leadership is a composite variable created from the multiplication of three variables: number of meetings principal had with supervisor at school, average length of visits with supervisor at school, and average percentage of visits spent focusing on instructional leadership. Overall number of minutes spent on instructional leadership remained similar over the course of the initiative.

Principals' expectations about their work with their supervisors changed. Principals' understanding of effective support changed during the PSI, as did the types of relationships they wanted to have with their supervisors. Principals came to expect a consistent relationship with their supervisors that included coaching, feedback, deeper professional development, and stronger supports for and expertise about instruction.

> **It's a tremendous change. I mean, we're talking goals, we're talking growth, and we're talking data. It's not just dropping by on a whim, it's purposeful.**
>
> ***—Principal (2018)***

Just as supervisors had to learn to accept their new role, some principals had to revise their expectations for what work with their supervisor looked like. The PSI required principals to accept more intensive, ongoing coaching and feedback from their supervisors. Some principals were not initially receptive to receiving coaching, more ongoing and in-depth feedback, and engaging with their peers through learning networks in new and more intensive ways. To overcome principals' hesitance, many supervisors and principals spoke of the importance of building relationships based in mutual trust and respect. This mutual trust and respect allow for honest and open conversations. Principals learned to see their supervisors as working in their best interests, rather than dropping in to catch

> **Now, whenever he comes, I know that is not going to be a simple 'okay great.' It's going to be a why, or did you think about, have you thought about, or could you think about. He's going to make me explore all my options and defend the choices that I've made. That's new.**
>
> ***—Principal (2018)***

them doing something wrong, which was often how principals described visits from supervisors prior to the PSI changes.

> **The authenticity of my relationship with my current boss, a foundation of trust affords us an opportunity to engage in both easy and difficult conversations, and because of the trust I have in her motivations, we can have the hard conversations and there's no issue with that.**
>
> ***—Principal (2018)***

Principals' perceptions of supervisor practices and effectiveness improved over time. The PSI was based on the idea that supervisors would become more effective in supporting principals by adopting key practices linked to the development of principal leadership and adult learning.

Principals rated supervisors on surveys in three key areas of supervisor effectiveness:[10]

1. **The quality of feedback they received from their supervisors on their formal evaluations.** Conducting evaluations and providing actionable feedback for improvement was a major emphasis of the PSI and part of supervisors' role in developing principals.
2. **How often their supervisors' engaged in 18 specific, focal practices that were emphasized by the PSI.** These practices included working with principals to jointly decide the principals' goals and monitoring principals' growth and change from one visit to the next.
3. **Overall effectiveness on 30 areas related to principal supervisors' work with principals.** These areas included providing principals with actionable feedback, helping principals improve their teachers' instruction, and helping principals assess their strengths and weaknesses.

Principal ratings of supervisors on these three key areas of supervisor effectiveness increased on average across the districts throughout the course of the initiative (Figure V.3). According to principals' reports, the quality of supervisors' evaluation feedback and supervisors' general effectiveness improved, and supervisors engaged more regularly in the PSI's focal practices.

[10] See Appendix A for information on individual items and scale construction.

Figure V.3. Principals' perceptions of supervisors' practices and effectiveness increased

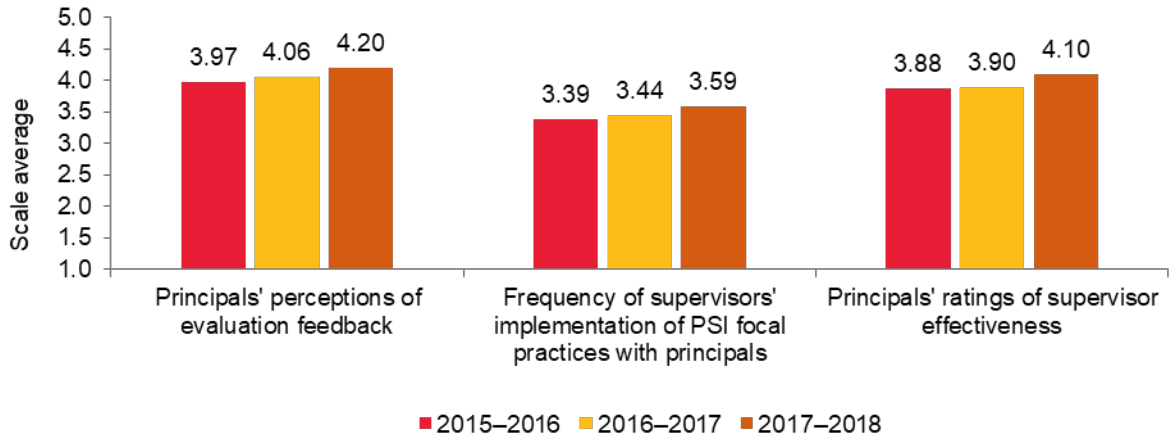


Figure reads: Principals rated the quality of their supervisors' evaluation feedback an average of 3.97 out of 5 in 2015–2016.

Source: Principal surveys, 2016 (*N* = 635), 2017 (*N* = 639), and 2018 (*N* = 606).

Note All scales range from 1 to 5. See Appendix A for information on scale creation. Changes across the three years were statistically significant at the 5 percent level for all measures.

C. Principals' perceptions of their leadership capacity and practices

Principals reported that they increased their leadership capacity. Central office staff in some districts and principals themselves indicated that increased support from supervisors facilitated a shift in principals' capacity to address instructional matters in their buildings. A Cleveland principal spoke of how work with her supervisor strengthened her ability to have focused conversations with teachers rooted in concrete examples from classroom walk-throughs. She felt the work with her supervisor improved her ability to assess the rigor and quality of instruction and developed her confidence in discussing their observations with teachers. A principal from Des Moines explained that his interactions with teachers were now purposeful, whereas before his classroom visits were typically just "popping in." This same principal noted he was leading a drive to improve student achievement and growth, "growing teachers," and creating an environment of learning for students and teachers. In Broward, district leaders attributed gains in student achievement, including moving some schools off the state's lower-performing schools list, to the PSI and to the close work between supervisors and principals.

> **Both the biggest change and probably the biggest accomplishment is there is an overwhelming sentiment that principals are far more instructional leaders and confident in their abilities to lead, support, and monitor instructional practices in their building as a result of the support and training and PD of their principal supervisor.**
>
> ***—District leader (2018)***

Some principals said the PSI changed their mindset and practice. Many principals also noted a shift in their own understanding of their roles as instructional leaders. Specifically, they felt more independent, valued, and focused on instruction. The strong focus on understanding and defining high-quality instruction in many of the districts meant that principals were able to

assess instructional quality and rigor. This in turn provided them with the ability to assess students' and teachers' needs. For example, many principals noted their classroom walk-throughs changed because of the shift in their own work with their supervisors; their roles became more focused on indicators of instructional quality and practice (for example, the quality of student discussions and participation). In the past, before the PSI, walk-throughs were commonly focused on big picture issues like classroom management or if the teacher projected confidence. The shift allowed principals to delve deeper into teacher practice and support because the observations were focused on specific areas of practice.

> **I feel like my time is better used keeping up with the curriculum and keeping up with what's going on…. Because it's very easy to get bogged down with operational things as a principal. [The way we work now] almost forces us to step out of that operational piece and make sure that we're on track with the curriculum piece.**
>
> ***—Principal (2018)***

Changes in the work between supervisors and principals also influenced how principals interacted with their school leadership teams.

> **I feel like because the focus has changed so much, I have sharpened my skills in the area of instruction, and my leadership meetings—my [meetings with my] assistant principal, my team leader meetings, my monthly leadership meetings with the support staff and team leaders, have definitely been more curriculum-focused rather than just an agenda with operational pieces.**
>
> ***—Principal (2018)***

D. Experiences varied across principals

Principals in every district reported positive experiences during the PSI. However, these experiences were not consistent across all districts. Although, on average, principals in every district reported improvements in the quality of their district's central office, variation in central office quality was evident throughout the initiative. Principals in one district, for example, consistently rated their central office more highly than principals in other districts, and principals in a second district consistently rated their central office more highly than principals in a third district. These differences highlight both the influence of district context in the ways the PSI unfolded as well as the variation across districts in principals' experiences over the course of the initiative.

Similarly, not every principal could connect the work they did with their supervisor to changes in their own practice. One-on-one interactions between supervisors and principals were a crucial component of principal supervision, and principal/supervisor relationships varied from principal to principal and supervisor to supervisor. For example, principals' descriptions of the helpfulness and value of their work with their supervisor varied, even as principals' assessments of their supervisor's effectiveness generally improved over time. Additionally, principals emphasized the importance of interpersonal factors in determining the success of their work with

their supervisor, including interpersonal compatibility, working style, trust, and confidence in their supervisor's level of expertise and familiarity with their school. Principals noted that learning from their supervisor became more difficult when these factors were absent.

A spotlight on principal evaluation and the Principal Supervisor Initiative

Principals had positive views of the evaluation and feedback process. In 2018, 100 percent of principal supervisors in the PSI districts reported that they were responsible for evaluating principals. While principal evaluation had always been a key part of the principal supervisor role, during the PSI the districts worked to better align principal evaluation with principal support. Although most districts did not change the evaluation systems they used over the course of the PSI, principals detected changes in *how* districts and supervisors implemented these systems.

There is little evidence of conflict between supervisors' dual roles as coach and evaluator. Supervisors worked to balance their dual role as both coach and evaluator. While this dual role has historically been viewed as a potential trigger for conflict, many personnel at all levels in PSI districts applauded the dual role. They felt that increased ongoing support allowed supervisors to gain the trust, firsthand knowledge of the principals' leadership, and understanding of school context needed to evaluate principals accurately and fairly.

> **I think sometimes [the dual role] can be difficult…. There are benefits to it too, that you're building that relationship. That person, if they're with you for a length of time, really gets to know you and can share with you honestly and openly what it is they feel you need to grow in, what it is you need. But I think that relationship, being able to build that relationship, helps with even the strongest way of evaluating a person.**
>
> ***—Principal (2018)***

Principal supervisors tried to be clear about when they were inhabiting one role or the other, even though shifting between the two was common during the same interaction. Owing to the interpersonal skills required to make these shifts, participants cited a trusting relationship between supervisor and principal as the most important aspect of successfully balancing the dual role. Without trust, principals could feel they needed to be cautious about what they revealed to a supervisor during the coaching portion of their interactions. A principal explained that hypothetically: "You might feel that trust isn't as strong. And you don't want them to use things that you've shared in your coaching as opportunities to reflect negatively in your evaluation." Despite this potential pitfall, most principals believed their supervisors navigated the dual role well.

Integrating formative and summative approaches to principal evaluation was important for supporting principals. Districts that grounded their principal evaluation systems in formative assessment—processes to provide ongoing, constructive feedback—created an evaluation culture where the ongoing work between principal and supervisor was tightly coupled with evaluation. For example, in Broward, Des Moines, and Long Beach, principal evaluation was viewed as a learning tool that offered principals continuous formative assessment and feedback. Principals in these districts

often described evaluation as a nonthreatening, reflective process that focused on leadership growth and aligned with their day-to-day work.

Improved supervisor implementation of current evaluation systems increased the integration of evaluation and support. For example, principals in Broward felt that evaluation quality improved because supervisors spent more time working directly with principals in their schools. This increased time together produced deeper conversations between principals and supervisors and allowed supervisors to ground their evaluation in deep knowledge of the school and the work of the principal.

> **We would've had to talk about our self-evaluation with our boss, and we would've had to have a mid-year and an end-of-year, already before [the PSI]]. But now you can spend a lot more time. You can go through the protocols and talk about it…. You could actually walk-through multiple classrooms and talk about what you're seeing in classrooms, and not just walk around the outside hallways for a few minutes and do just like a drive-by…. The depth was definitely better. The substance was definitely better.**
>
> ***—Principal (2018)***

Not all districts were able to integrate formative assessment and ongoing supervisor support. In some districts, evaluation remained highly procedural and entirely summative, providing little opportunity for principals to reflect and use outcomes for future development. This type of evaluation created a sense among principals that evaluation was a compliance-driven activity all about "checking boxes."

Principals' perceptions of the quality of written and oral feedback from evaluations increased over the years of the PSI. Principals' overall perception of the quality of their district's principal evaluation system rose steadily. In 2016, 53 percent of principals agreed or strongly agreed that their evaluation provided actionable feedback they could use to improve their leadership; by 2018, this number had risen to 62 percent (Figure V.4).

Figure V.4. The percentage of principals reporting that the evaluation system provided more actionable feedback increased

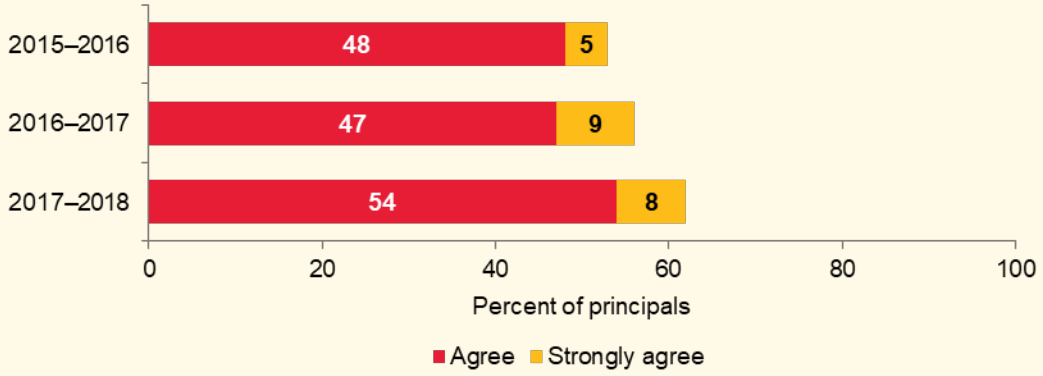


Figure reads: Forty-eight percent of principals agreed and 5 percent strongly agreed that the principal evaluation system provided actionable feedback to improve their leadership in 2015–2016.

Source: Principal surveys, 2016 (*N* = 635); 2017 (*N* = 639); and 2018 (*N* = 606).

Note: The survey prompt read, "Please indicate to what extent you agree or disagree with the following statements, based on your experiences this school year." Changes across the three years are statistically significant at the 5 percent level.

Principal ratings of the quality of the oral feedback they received from their supervisor in the evaluation process also improved over time (see Figure V.5). Principals rated the quality of their written feedback similarly.

Figure V.5. Principals' perceptions of the quality of oral feedback received from evaluations improved

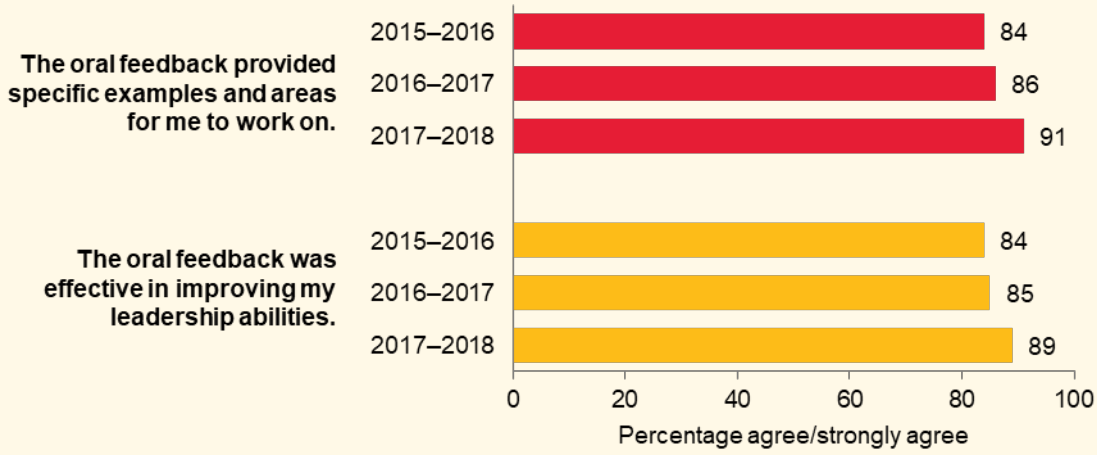


Figure reads: Eighty-four percent of principals agreed or strongly agreed that the oral feedback they received from their evaluations provided specific examples and areas to work on in 2015–2016.

Source: Principal surveys, 2016 (*N* = 635), 2017 (*N* = 639), and 2018 (*N* = 606).

Note: The survey question read, "Thinking about the oral feedback you received on your evaluation from your principal supervisor, please indicate to what extent you agree or disagree with the following statements." Changes across the three years are statistically significant at the 5 percent level for both items.

Challenges to improving evaluation

Several districts faced ongoing challenges to implementing an effective evaluation process:

Supervisors sometimes lacked consistency in their evaluation ratings of principals. Principals spoke of a lack of calibration or consistency from one supervisor to the next. When principals perceived that different supervisors gave different ratings on the same evaluation system, they felt that their evaluations were invalid and unfair. Districts were keenly aware of this pitfall, and they spent varying degrees of effort working with principal supervisors to calibrate how they rated principals on the evaluation rubric.

> **I would say the heaviest lift has been the consistency of principal supervisors as they evaluate schools … the consistency of just some common language and things of that nature is really a heavy lift.**
>
> ***—Principal supervisor (2018)***

Not all supervisors spent enough time in schools to accurately assess principals' performance. When supervisors were present in schools on a regular basis, and were able to see principals' work consistently, principals felt that supervisors better understood the principal's strengths and weaknesses. This firsthand knowledge was integrated into the evaluation process and added legitimacy to it. However, not all supervisors spent enough time in schools. When supervisors had little contact with principals, evaluation results were deemed not useful or illegitimate by principals.

VI. THE PSI'S EFFECTS ON TEACHERS' PERCEPTIONS OF PRINCIPALS' PERFORMANCE

According to the PSI's theory of change, enhanced support for principals was intended to improve their performance in terms of their instructional leadership skills and practices. However, our analysis of the VAL-ED data did not find positive effects of the PSI on teachers' perceptions of principals' performance.

A. The PSI's average effects on teachers' perceptions of principals' performance

The PSI did not affect teachers' perceptions of principals' performance. Despite the positive experiences principals reported with the PSI, the initiative did not have meaningful effects on teachers' perceptions of principals' performance. (As noted in Chapter II, we considered a meaningful effect to be equivalent to changing an average principal into a high-performing principal. We defined average as scoring in the 50th percentile on the baseline VAL-ED among all principals with VAL-ED scores in the 2013–2014 school year and high-performing as scoring in the 75th percentile.) On average across all the schools in the sample, differences in teachers' perceptions of principals' performance between PSI and matched comparison schools were small and not statistically significant (Figure VI.1). Estimated effects in all four years of the initiative were well below our definition of a meaningful effect. The estimated effects were slightly positive in some years and slightly negative in others, helping to rule out the possibility that the program had consistently positive or negative effects that were simply too small to detect.[11]

[11] We conducted nine sensitivity analyses that varied based on aspects of the estimation approach and sample (see Appendix B for details). Effects estimated from these analyses were also small and not statistically significant (Appendix Table B1).

Figure VI.1. Differences in teacher-reported VAL-ED scores between PSI and matched comparison schools by school year were not statistically significant

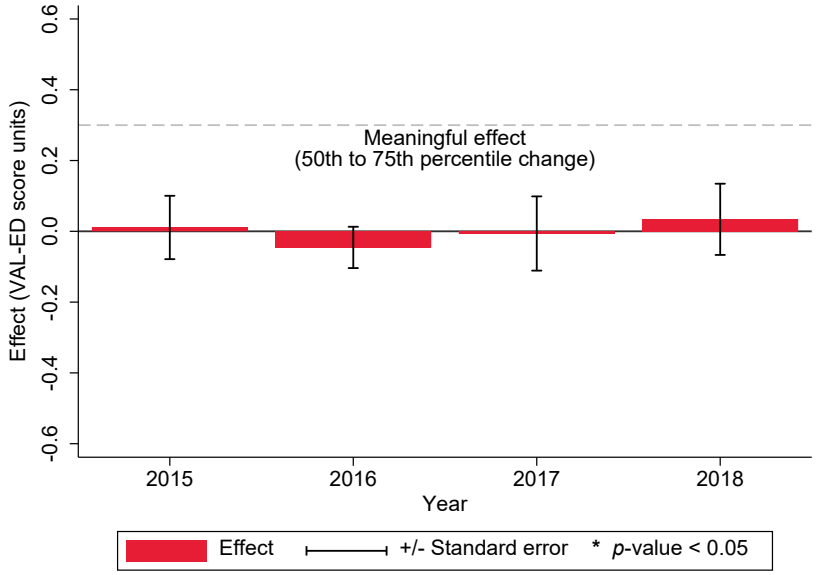


Figure reads: The PSI did not have a statistically significant effect on teachers' perceptions of principals' performance in any of the four years of the initiative.

Sources: VAL-ED survey data (2013–2014 to 2017–2018 school years), Common Core of Data (2013–2014 school year), and EdFacts data (2011–2012 to 2013–2014 school years).

Note: The sample excludes districts that are in The Wallace Foundation's Principal Pipeline Initiative and two exemplar districts that received smaller PSI grants but are not part of the evaluation. Because all schools in PSI districts were classified as either urban or suburban, we also restrict the comparison schools to those classified as urban or suburban. We defined a meaningful effect to be equivalent to changing an average principal into a high-performing principal. We defined average as scoring in the 50th percentile on the baseline VAL-ED among all principals with VAL-ED scores in the 2013–2014 school year and high-performing as scoring in the 75th percentile.

We also did not find any statistically significant effects when analyzing elementary and secondary schools separately (Figure VI.2). The estimated effect for elementary schools during 2017–2018 was relatively large. However, we cannot rule out that it arose due to chance given the large standard error.

As noted in Chapter II, we examined the PSI's effects on teachers' perceptions of principals' performance, based on teacher-reported VAL-ED scores, due to concerns that supervisors' and principals' reports of principals' performance could be biased. Such bias might arise because the PSI might have affected not only the true performance of principals, but also *how* supervisors rated principals and how principals rated themselves on the VAL-ED. For instance, the initiative could have led principal supervisors to hold principals to a higher standard, and principals to hold themselves to a higher standard. If so, it might not be appropriate to compare supervisor- or principal-reported VAL-ED scores (or aggregate scores based on supervisor, principal, and teacher reports) in the PSI schools with those in comparison schools. Despite these concerns, exploratory analyses show that the PSI did not affect principals' performance as measured by principal- and supervisor-reported VAL-ED scores or aggregate scores (Table B.2), consistent with our main results.

Figure VI.2. Differences in teacher-reported VAL-ED scores between PSI and matched comparison schools were not statistically significant at the elementary or secondary school level

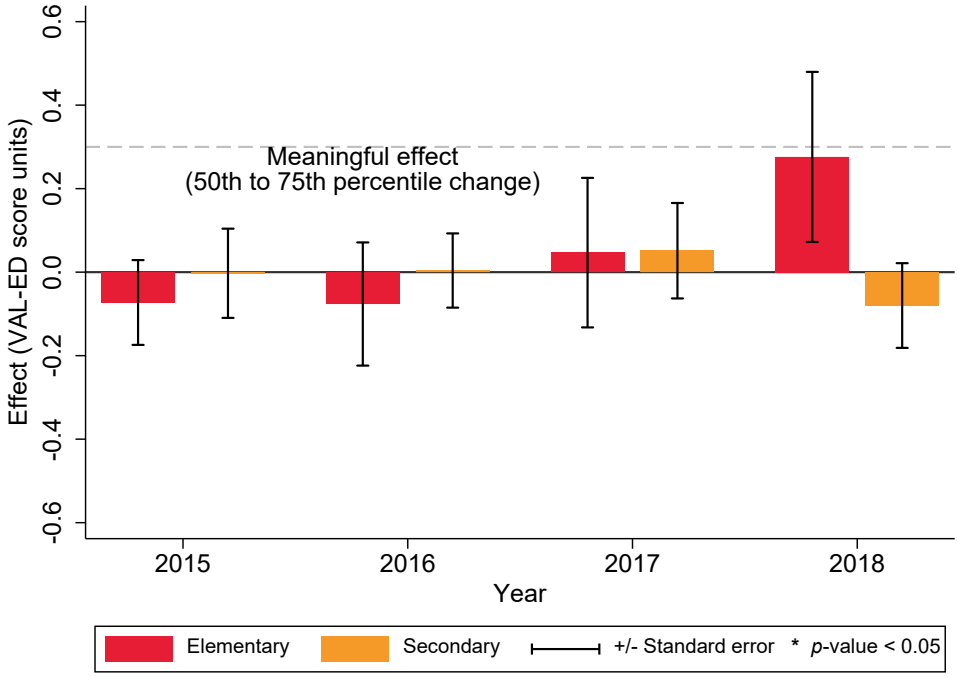


Figure reads: The PSI did not have a statistically significant effect on teachers' perceptions of principals' performance at the elementary or secondary level in any of the four years of the initiative.

Sources: VAL-ED survey data (2013–2014 to 2017–2018 school years), Common Core of Data (2013–2014 school year), and EdFacts data (2011–2012 to 2013–2014 school years).

Note: The sample excludes districts that are in The Wallace Foundation's Principal Pipeline Initiative and two exemplar districts that received smaller PSI grants but are not part of the evaluation. Because all schools in PSI districts were classified as either urban or suburban, we also restrict the comparison schools to those classified as urban or suburban. We defined elementary schools as those for which all students are in grades six or lower and secondary schools as all other schools. We defined a meaningful effect to be equivalent to changing an average principal into a high-performing principal. We defined average as scoring in the 50th percentile on the baseline VAL-ED among all principals with VAL-ED scores in the 2013–2014 school year and high-performing as scoring in the 75th percentile.

B. Variation in teachers' perceptions of principals' performance and the PSI's effects across supervisors and districts

Although the PSI had no measurable effects on teachers' perceptions of principals' performance on average across the six districts in the initiative, it is useful to examine how teachers' perceptions of principals' performance varied across districts and supervisors and how the PSI's effects varied across districts and supervisors.

- *Examining variation in teachers' perceptions of principals' performance* can shed light on the extent to which districts and principal supervisors influenced principals' performance, irrespective of any effects of the PSI. As discussed in Chapter II, we measured teachers'

perceptions of principals' performance using teachers' reports on principals' performance from the VAL-ED.

- *Examining variation in the PSI's effects on teachers' perceptions of principals' performance* can shed light on the extent to which the PSI's effects differed across individual districts and principal supervisors. We measured the PSI's effects on teachers' perceptions of principals' performance by comparing teachers' reports on principals' performance from the VAL-ED across PSI schools and a similar set of matched comparison schools, as described in Chapter II.

We found that principals' performance, as well as the PSI's effects on teachers' perceptions of principals' performance, did not vary systematically across districts or supervisors. Although teachers' perceptions of principals' performance and the PSI's effects varied across schools, individual districts and supervisors accounted for little of this variation. This suggests that districts' differing approaches to implementing the PSI and the behaviors of individual supervisors were not driving the differences in effects across principals.

Districts and supervisors did not have much influence on teachers' perceptions of principals' performance. Although teachers' perceptions of principals' performance varied across schools, districts accounted for at most 10 percent of the variation in these perceptions in the PSI districts, and supervisors accounted for none of the variation (Figure V.8), after adjusting for potential measurement error across the four years of PSI implementation. Similarly, in comparison districts, districts accounted for only 17 percent of the variation in teachers' perceptions of principals' performance, on average. These results suggest that the PSI districts and supervisors, as currently operating, do not have large effects on teachers' perceptions of principals' performance.

Similarly, districts and supervisors had limited influence on the PSI's effects on teachers' perceptions of principals' performance. After adjusting for potential measurement error in the PSI's effects across the four years of implementation, districts and supervisors each accounted for at most 3 percent of the variation in PSI effects (Figure VI.3). The remaining 94 percent of the variation in PSI's effects on teachers' perceptions of principals' performance was due to individual differences between schools.

Figure VI.3. Districts and supervisors accounted for little of the variation in teachers' perceptions of principals' performance and PSI effects

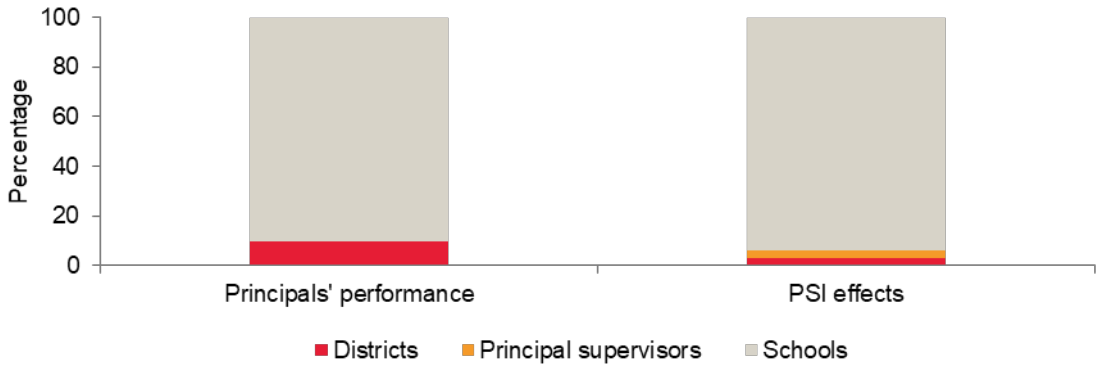


Figure reads: Districts accounted for 10 percent of the variation in teachers' perceptions of principals' performance and principal supervisors accounted for none of the variation.

Sources: VAL-ED survey data and estimated PSI effects from 2014–2015, 2015–2016, 2016–2017, and 2017–2018.

Note: The total proportion of variation in teachers' perceptions of principals' performance and PSI effects removes the proportion of variation across years within each school. We assume that the variation across year is measurement error. More details on the calculations are provided in Appendix A.

Another way to illustrate the variation in teachers' perceptions of principals' performance is through comparison of the distribution of VAL-ED scores across districts. As shown in Figure VI.4, PSI districts had similar, mostly overlapping distributions of VAL-ED scores in 2018, with averages that differed very little. In other words, knowing which district a school is in provides very little information about how well a principal performs (in the eyes of teachers). Each district had high-performing principals and low-performing principals, with few clear differences in the proportions of high- and low-performers. This reinforces the findings that the PSI districts accounted for a small portion of the variation in teachers' perceptions of principals' performance.

Figure VI.4. Distributions of teachers' perceptions of principals' performance were similar across PSI districts

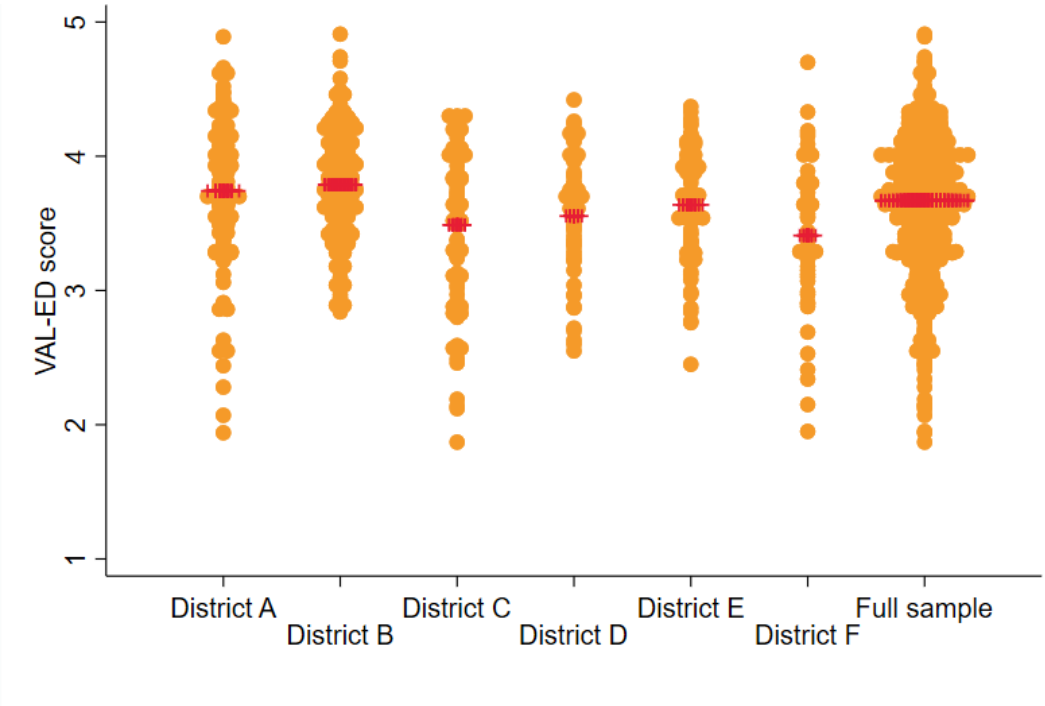


Figure reads: PSI districts had similar, mostly overlapping distributions of teachers' perceptions of principals' performance in 2018.

Source: VAL-ED survey data collected in 2017–2018.

Note: The yellow dots are VAL-ED score for each school and the red bars are the mean VAL-ED scores in the district.

C. Possible explanations for the PSI's lack of effects on teachers' perceptions of principals' performance

Although our analysis does not tell us *why* the PSI did not affect teachers' perceptions of principals' performance, several factors could have contributed to the lack of effects.

- **The study examined only one measure of principals' performance.** Although the VAL-ED is a valid and reliable measure of teachers' perceptions of principals' performance, it may not capture all dimensions of principals' performance. The PSI could have affected other aspects of performance not captured by VAL-ED, or not apparent to teachers, such as principals' abilities to hire or retain good teachers or to positively affect student achievement.

- **It may take more time for changes at district and supervisor levels to affect teachers' perceptions of principals' performance.** The timeframe for the study may have been too short to detect changes in teachers' perceptions of principals' performance. For example, it may have taken time for supervisors to adjust to their new roles and for principals to refine their leadership strategies and practices based on their interactions with their supervisors. Additionally, teachers may not yet have noticed changes in principals' performance or have

based their assessment on their history of experiences with their principal, including those that predated the PSI.

- **It may be difficult to influence teachers' perceptions of principals' performance through an initiative focused solely on districts and supervisors.** If districts and supervisors have limited influence on their principals' performance, then initiatives like the PSI that are focused at the district and supervisor level but do not interact directly with principals might have limited influence on principals' performance. This possibility is supported by our findings that districts and supervisors accounted for little of the variation in teachers' perceptions of principals' performance. It is possible that more direct engagement with and buy-in from principals could have led to more positive effects. The Wallace Foundation's Principal Pipeline Initiative, for example, produced measurable impacts using an approach that involved directly intervening with principals themselves (Gates et al. 2019). A broader initiative that directly touched districts, principal supervisors, and principals may have had more positive effects.

- **Districts may have required more specific guidance on implementation.** The PSI gave districts considerable opportunity to adapt and vary the initiative to meet their own needs. Although this approach provided flexibility to the districts, the initiative might have been more effective if districts had implemented it in a more uniform way or supervisors had engaged in a uniform set of approaches.

- **Some districts may have lacked the capacity for strong implementation.** As discussed in Chapter IV, some districts faced challenges related to PSI implementation, particularly in the later years of the initiative. Districts may have required additional support and strong district leadership to sustain the changes made in the initial years of the initiative. In addition, accountability pressures and local contexts led some districts to adopt multiple initiatives, which may have diverted their focus away from the goals of the PSI.

- **Other districts may have made similar changes to principal supervision during the PSI.** The estimates of the PSI's effects are based on a comparison of principals' performance in the PSI schools and a similar set of non-PSI schools. However, if the comparison schools' districts made similar changes to their central office structures and the principal supervisor role during the same period, we might not expect to see any effects of the PSI. In fact, a comparison of the principal supervisor role in the PSI districts and a sample of other urban districts found that many other urban districts had made similar changes to those in the PSI districts. We discuss these findings further in Chapter VIII.

Summary. Principals reported positive experiences with the PSI, including more positive perceptions of the support they received from the central office, their work with their supervisors, and their supervisors' effectiveness. They also reported an increase in their leadership capacity, and some reported a shift in their understanding of their role as instructional leaders. However, experiences varied across principals and districts, and not every principal could connect the work they did with their supervisor to changes in their own practices.

The PSI did not affect teachers' perceptions of principals' performance as measured by the VAL-ED. Furthermore, districts and supervisors accounted for little of the variation in teachers' perceptions of principals' performance.

VII. THE RELATIONSHIP BETWEEN COMPONENTS OF THE PSI AND ITS EFFECTS ON TEACHERS' PERCEPTIONS OF PRINCIPALS' PERFORMANCE

Although our analysis of VAL-ED data found that the PSI had no effects on teachers' perceptions of principals' performance on average across all the schools in the initiative, effects varied considerably across individual schools. In addition, as discussed in Chapter III, the PSI was not highly prescriptive but instead encouraged participating districts to implement the components in ways that aligned with their particular contexts and priorities. Accordingly, districts varied in their specific approaches to implementing the initiative, and supervisors also varied in the amount of time they spent with principals and their effectiveness in supporting principals, leading to considerable variation in the experiences of individual principals and schools. The variation in PSI implementation and effects across schools suggests that differences in how individual districts and supervisors implemented the PSI could relate to differences in the PSI's effects across schools.

To further explore the variation in the PSI's effects on teachers' perceptions of principals' performance, we examined how specific components of the PSI and principal supervisors' time spent on instructional leadership related to the PSI's effects on principals' performance. We also examined how supervisors' practices and effectiveness, as perceived by principals, related to the initiative's effects. The findings from these correlational analyses, summarized in Table VII.1 and described further below, were not sensitive to the choice of estimation model or sample (see Appendix B).

Table VII.1. Summary of associations between implementation factors and PSI effects on teachers' perceptions of principals' performance

	Statistically significant at the 5 level	Not statistically significant at the 5 level
Small positive association with PSI effects on teachers' perceptions of principals' performance	• Supervisors' span of control • Percentage of supervisor's time spent on instructional leadership in past 3 months • Principals' ratings of supervisor's effectiveness • Supervisor's implementation of PSI's focal practices • Principals' perceptions of evaluation feedback	• Quality of central office support according to principals • Supervisor's time spent on instructional leadership in past 3 months, as reported by principals (minutes)
Small negative association with PSI effects on teachers' perceptions of principals' performance	• n.a.	• Supervisor training quality • Alignment of supervisor training with PSI goals • Quality of central office support according to supervisors

Sources: Principal and supervisor survey data from 2016, 2017, and 2018 and VAL-ED survey data from 2018.

Notes: Implementation measures are averaged across the three years in which principals and supervisors were surveyed: 2015–2016, 2016–2017, and 2017–2018.

For all regression models, the dependent variable is the PSI effects on principals' VAL-ED scores in 2018.

n.a. = not applicable. There were no negative, statistically significant correlations.

A. Relationship between implementation of PSI components and the PSI's effects on teachers' perceptions of principals' performance

Measures of districts' implementation of three PSI components had little to no relationship with the PSI effects on teachers' perceptions of principals' performance. We examined measures of PSI implementation for the three PSI components that could vary across supervisors within districts and were hypothesized to influence supervisor performance: (1) principal supervisors' span of control, (2) the quality of supervisor training and its alignment with PSI goals, and (3) the quality of support provided by the central office to supervisors and principals. (Chapter II and Appendix A provide more details on the construction of these measures.)

None of these measures of PSI implementation were meaningfully related to the PSI's effects on teachers' perceptions of principals' performance. The associations between measures of supervisor training and central office support and PSI effects on teachers' perceptions of principals' performance were statistically indistinguishable from zero and close to zero in magnitude (Appendix Table B.6). The measure of supervisor span of control had a statistically significant positive association with PSI's effects, indicating that, on average, PSI effects were *larger* among supervisors overseeing *more* principals than among supervisors overseeing fewer principals. This is a counterintuitive relationship given that reduced span of control was intended to improve the quality of support supervisors could give principals. However, the association between supervisors' span of control and the PSI's effects on teachers' perceptions of principals' performance was so small that we do not consider it to be meaningfully related to the PSI's effects. To assess the magnitude of the association, we considered how a change in span of control would affect the performance of a low-performing principal (one at the lowest quartile of performance). The estimated coefficient of 0.02 means that decreasing a supervisor's span of control by two principals is associated with a 0.04 decrease in PSI effects, which would move a principal from the 25th percentile to the 23rd percentile of performance on the VAL-ED, among principals in PSI districts in 2015 (Figure VII.1).

Figure VII.1. The change in teachers' perceptions of principals' performance associated with decreasing a supervisor's span of control by two principals was negligible

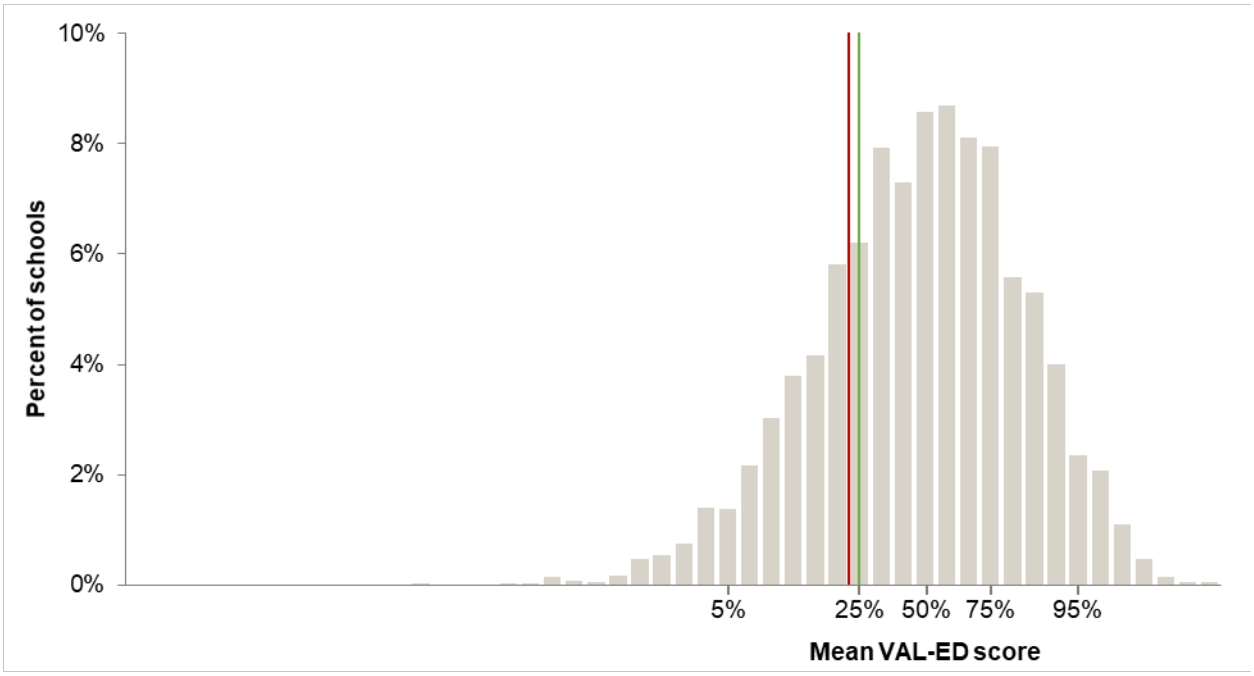


Figure reads: A reduction in supervisor's span of control by two principals is associated with moving a principal from the 25th percentile to the 23rd percentile of performance on the VAL-ED, among principals in PSI districts in 2014–2015.

Source: VAL-ED baseline survey data (2013–2014), estimated PSI effects from 2014–2015, 2015–2016, 2016–2017, and 2017–2018, and supervisor survey data (2015–2016, 2016–2017, and 2017–2018).

B. Relationship between supervisors' time spent on instructional leadership and the PSI's effects on teachers' perceptions of principals' performance

Principal supervisors' time spent on instructional leadership activities with each principal also had little to no relation with the PSI's effects on teachers' perceptions of principals' performance. We measured principal supervisors' time spent on instructional leadership activities in two ways. We constructed the first measure as the product of three principal survey items: (1) frequency of meetings with the principal supervisor, (2) length of meetings with the principal supervisor, and (3) percent of time spent with the principal supervisor on instructional leadership activities. The estimated association between this composite measure of supervisors' average total time spent on instructional leadership activities, as reported by principals, and the PSI's effects on teachers' perceptions of principals' performance was statistically indistinguishable from zero and close to zero in magnitude (Figure VII.2).

To account for the possibility of measurement error in our measure of supervisor time spent on instructional leadership, we examined a second measure of supervisor time on instructional leadership. Respondent interpretations of the survey questions about length and frequency of principal meetings with supervisors may have varied widely. Our alternate measure of supervisors' time spent on instructional leadership was based on a single principal survey item: principals' reports of the percentage of time spent working with their supervisor that was spent on instructional leadership activities. The association between this measure of supervisors' time spent on instructional leadership and the PSI's effects on teachers' perceptions of principals' performance was statistically significant but close to zero. For example, the estimated coefficient of 0.002 suggests that a 20 percentage point (one standard deviation) increase in the percentage of time a supervisor spent on instructional leadership is associated with moving a principal at the 25th percentile just two percentage points higher on the distribution of performance on the VAL-ED, to the 27th percentile of performance. The very small magnitude suggests that the percentage of supervisors' time spent on instructional leadership was also not meaningfully related to the PSI's effects.

Figure VII.2. The change in teachers' perceptions of principals' performance associated with a 20 percentage point increase in supervisors' time spent on instructional leadership was small

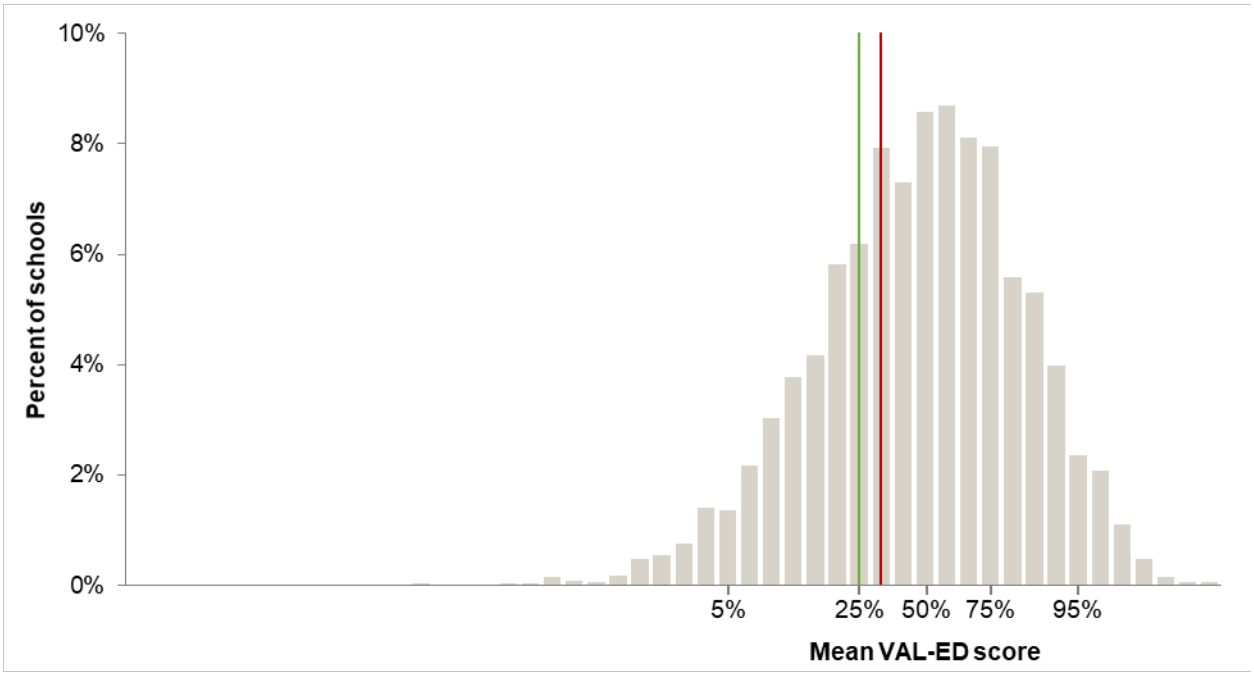


Figure reads: A 20 percentage point increase in the percentage of time a supervisor spent on instructional leadership is associated with moving a principal from the 25th percentile to the 27th percentile of performance on the VAL-ED, among principals in PSI districts in 2014–2015.

Source: VAL-ED baseline survey data (2013–2014), estimated PSI effects from 2014–2015, 2015–2016, 2016–2017, and 2017–2018, and supervisor survey data (2015–2016, 2016–2017, and 2017–2018).

C. Relationship between principals' perceptions of supervisor effectiveness and the PSI's effects on teachers' perceptions of principals' performance

Principal supervisors' effectiveness, as perceived by principals, had a small positive relationship with the PSI's effects on teachers' perceptions of principals' performance. We examined three measures of effectiveness, all constructed as composites of principal survey items and all on a 1–5 scale: (1) perceived overall effectiveness, (2) perceived quality of evaluation, and (3) reported implementation of practices aligned with the PSI. Each of the three measures of supervisors' effectiveness we examined had a positive and statistically significant association with PSI's effects. The PSI's effects on teachers' perceptions of principals' performance were higher for principals who worked with supervisors who provided high-quality feedback, aligned their supervisory practices with the PSI, and were regarded as effective overall.

However, the associations were all small. The estimated coefficients for supervisor effectiveness and alignment of supervisor practices with the PSI were both 0.11, meaning that a 1-point increase in principal-rated supervisor effectiveness (on a 1–5 scale) corresponds to an increase of .11 in the PSI's effects on teachers' perceptions of principals' performance on the VAL-ED. To better contextualize this magnitude, we calculated the change in teachers'

perceptions of principals' performance associated with incremental changes in PSI effects. A coefficient of 0.11 means that increasing a supervisor's perceived effectiveness (or alignment of practices with the PSI) by as much as two standard deviations (a very large change in perceived effectiveness, equal to the difference between a principal supervisor at the 25th percentile of effectiveness and one at the 95th percentile) is predicted to move a principal at the 25th percentile of the distribution of performance on the VAL-ED to the 33rd percentile (Figure VII.3). Similarly, the estimated coefficient of 0.9 for principal-reported quality of supervisor feedback predicts only a small improvement in teachers' perceptions of principals' performance (from the 25th percentile to the 32nd percentile) per a two standard deviation increase of the quality of supervisor feedback.

Although we cannot draw conclusions about the causal relationship between perceived supervisor effectiveness and PSI effects, these findings suggest that supervisors' approaches to their work with principals during the PSI were positively related to the PSI's effects on teachers' perceptions of principals' performance. Consistent with the finding that supervisors and districts have limited influence on teachers' perceptions of principals' performance and the PSI's effects, however, the small magnitudes of the relationships between supervisor effectiveness measures and PSI effects indicate that changes in supervisor practices during the PSI had limited capacity to improve teachers' perceptions of principals' performance.

Figure VII.3. The change in teachers' perceptions of principals' performance associated with a two standard deviation increase in principal perceptions of supervisor effectiveness (overall) was small

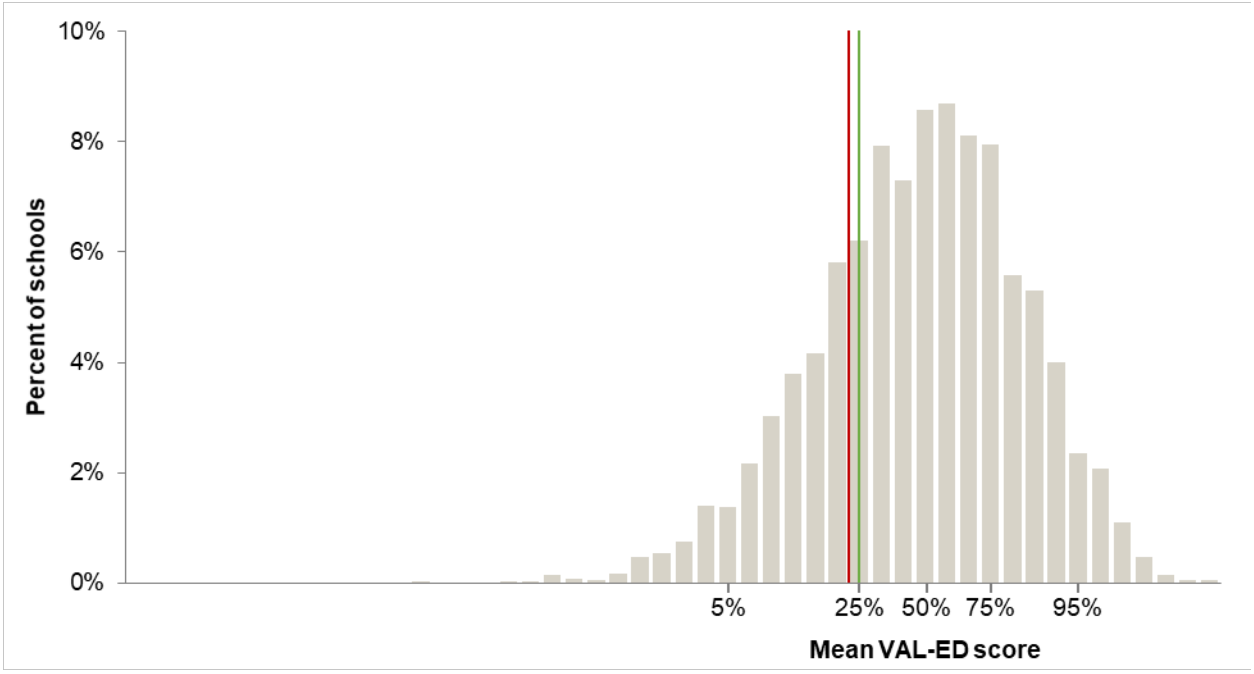


Figure reads: Increasing a supervisor's perceived effectiveness by two standard deviations is associated with moving a principal from the 25th percentile to the 33rd percentile of performance on the VAL-ED, among principals in PSI districts in 2014–2015.

Source: VAL-ED baseline survey data (2013–2014), estimated PSI effects from 2014–2015, 2015–2016, 2016–2017, and 2017–2018, and supervisor survey data (2015–2016, 2016–2017, and 2017–2018).

Summary. Although both approaches to PSI implementation and PSI effects varied across districts, the implementation factors we examined appeared largely unrelated to the PSI's effects on teachers' perceptions of principals' performance. Measures of supervisor practices and effectiveness and time spent on instructional leadership activities were all significantly positively associated with PSI effects, indicating that effects were more positive at schools assigned to supervisors who were perceived as more effective, more aligned with the PSI in their supervision practices, and more focused on instructional leadership. However, the magnitudes of the associations between all these implementation factors and the PSI's effects were small. This suggests that there may be other important factors influencing the PSI's effects that are not captured by the current theory of change as measured in this study.

VIII. A COMPARISON OF THE PRINCIPAL SUPERVISOR ROLE IN PSI DISTRICTS AND OTHER URBAN DISTRICTS

As the PSI districts strove to define and support the new principal supervisor role, other urban districts across the country also focused on improving the quality of supervision and support for principals (Cochran et al. 2020). National conferences focused on principal supervision, new principal supervision standards, and other local efforts might have prompted changes to the principal supervisor role in districts throughout the country. For example, in 2015, the Council of Chief State School Officers released its Model Principal Supervisor Standards (Council of Chief State School Officers 2015). These standards set new expectations for the principal supervisor role, focused on supporting principals as instructional leaders. A comparison of principal supervision in the PSI districts and other urban districts in the final year of the PSI shows several key differences, along with some important similarities. This comparison can provide important context for the changes to the principal supervisor role in the PSI districts. It can also shed light on a potential reason for a lack of effects on teachers' perceptions of principals' performance in the PSI districts.

A. Principal supervisors' span of control

Principal supervisors' span of control was lower in PSI districts than in other urban districts. The PSI's emphasis on reduced span of control was consistent with a broader trend toward reduced span of control in urban districts across the country (Cochran et al. 2020). Nonetheless, by the final year of the PSI, average span of control was lower in the PSI districts than the other urban districts in the survey sample (Figure VIII.1). Supervisors in PSI districts oversaw 13 principals on average, compared with an average of 16 principals for supervisors in the other districts. The range of span of control also differed between the PSI districts and the other urban districts. The span of control in PSI districts ranged from 7 to 19, compared with a range of 2 to 50 principals in the other urban districts.

Figure VIII.1. Principal supervisors' span of control was lower in PSI districts than in other urban districts

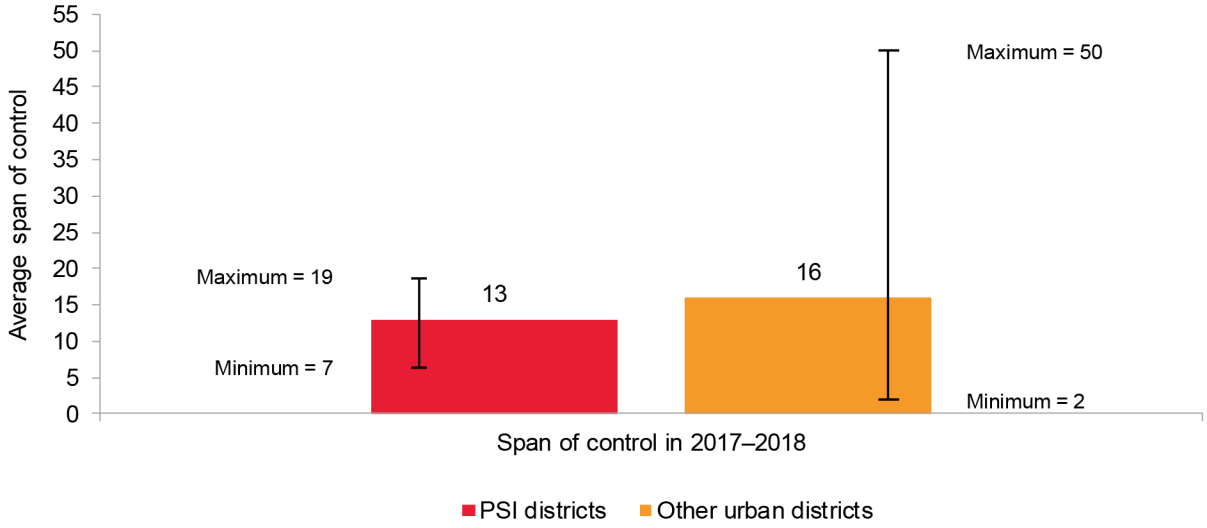


Figure reads: The average supervisor in a PSI district oversaw 13 principals.

Source: Wallace Foundation National Survey of Principal Supervisors, 2018 (supervisors in PSI districts, *N* = 50; supervisors in other urban districts, *N* = 293).
Note: Survey question read, "How many principals do you currently supervise?" Differences between supervisors in PSI and other urban districts are significant at the 5 percent level.

B. Training and support for principal supervisors

Principal supervisors in PSI districts were more likely to participate in role-specific training than those in other urban districts. Consistent with the PSI's emphasis on providing specific, targeted training to principal supervisors, supervisors in PSI districts were more likely than those in other urban districts to report participating in professional development related to their role as a principal supervisor (Figure VIII.2). In PSI districts, 80 percent of supervisors reported that they participated in training specific to their role in the 2017–2018 school year, compared with only 62 percent of supervisors in other urban districts.

Figure VIII.2. Principal supervisors in PSI districts were more likely to participate in role-specific training than those in other urban districts

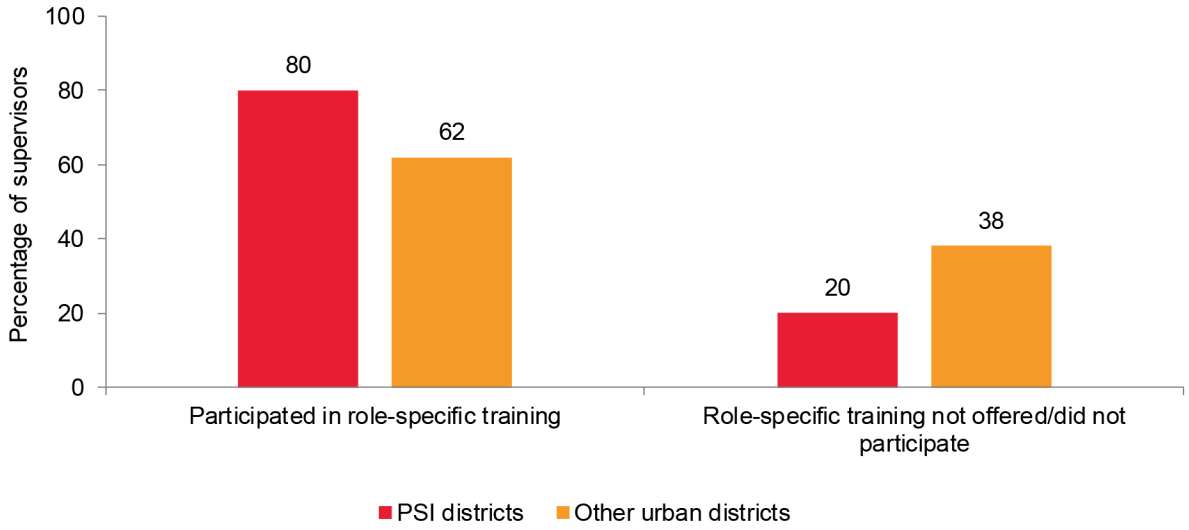


Figure reads: Eighty percent of principal supervisors in PSI districts participated in training specific to their role.

Source: Wallace Foundation National Survey of Principal Supervisors, 2018 (supervisors in PSI districts, *N* = 50; supervisors in other urban districts, *N* = 293).

Note: The survey question read, "During the 2017–2018 school year and the summer that preceded it, did you participate in district-sponsored training or professional development pertaining to your specific role as a principal supervisor?" Differences between supervisors in PSI and other urban districts are significant at the 5 percent level.

Supervisors in PSI districts rated the quality of their role-specific training more highly than supervisors in other urban districts. Supervisors in PSI districts were not only more likely to receive role-specific training, they also rated the quality of this training more highly than supervisors in other urban districts. Across multiple indicators of training quality, supervisors in PSI districts rated their training 3.7 points out of 5.0, whereas supervisors in other urban districts rated their training 3.5 points. (See Appendix A, Table A.3 for the full set of survey items included in this scale.)

Supervisors in PSI districts were more likely to report that their district offered training for new and aspiring principal supervisors than those in other urban districts. Consistent with the PSI's emphasis on succession planning, principal supervisors in PSI districts were significantly more likely to report that their district offered programs aimed at recruiting and training aspiring principal supervisors (Figure VIII.3). Fifty-nine percent of principal supervisors in PSI districts reported that their districts offered such programs, compared with 19 percent of supervisors in other urban districts. Supervisors in PSI districts were also more likely to report that their district offered mentoring and induction programs for new supervisors (49 percent in PSI districts compared with 22 percent in other urban districts). These differences are consistent with efforts in some of the PSI districts efforts to improve their identification and training of future principal supervisors (Chapter IV).

Figure VIII.3. Supervisors in PSI districts were more likely to report that their district offered programs for new and aspiring principal supervisors than those in other urban districts

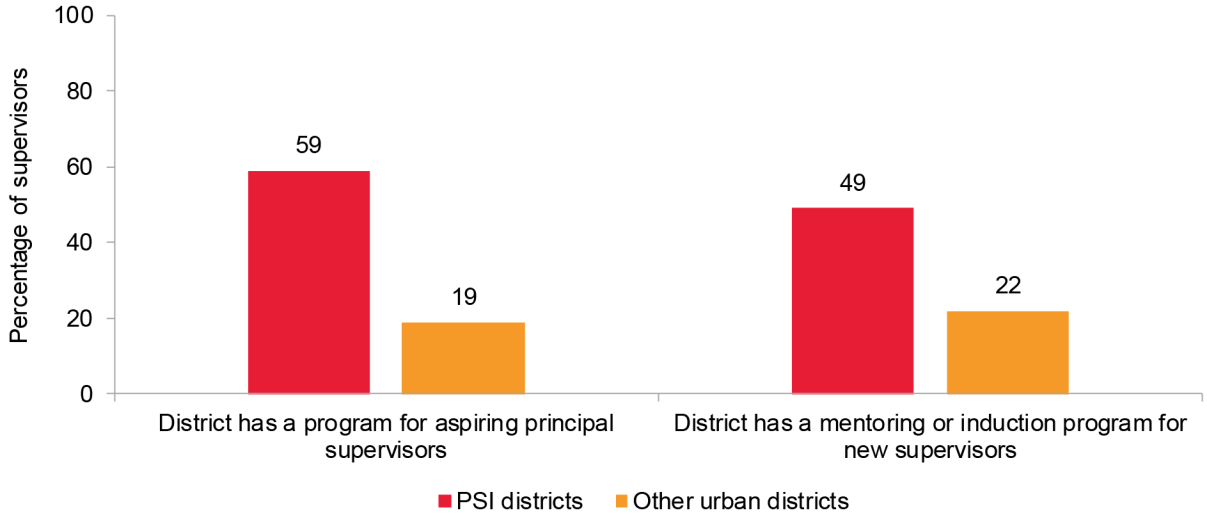


Figure reads: Fifty-nine percent of supervisors in PSI districts reported that their district has a program for aspiring principal supervisors.

Source: Wallace Foundation National Survey of Principal Supervisors, 2018 (supervisors in PSI districts, *N* = 50; supervisors in other urban districts, *N* = 293).

Note: The survey questions read, "Does your district have a program in place to identify and prepare aspiring principal supervisors?" and "Does your district have a mentoring or induction program for new principal supervisors?" Supervisors indicated yes or no. Differences between supervisors in PSI and other urban districts are significant at the 5 percent level for both items.

C. Principal supervisors' perceptions of central office support

Across both PSI and other urban districts, most principal supervisors agreed that improving teaching and learning was a key focus of the central office. However, most supervisors in both sets of districts disagreed that the central office was structured and focused in ways that supported their new roles (Figure VIII.4). For example, only 27 percent of supervisors in PSI districts and 36 percent of supervisors in other urban districts agreed that the central office understood their work. Similar proportions of supervisors in both sets of districts agreed that the central office facilitated supervisors' work with principals and was organized to support principals. Fewer than half of supervisors in both sets of districts agreed that central office meetings were scheduled so they could maximize their time in schools and that these meetings were useful.

Supervisors in PSI districts were more likely than those in other urban districts to be involved in deploying instructional support staff to schools. The PSI districts instituted new approaches to coordinating and deploying school supports. For example, some districts created central office support teams, with dedicated representatives from each central office department assigned to work directly with supervisors and principals; other districts created liaison structures, in which each supervisor was assigned as a liaison to another central office department (Goldring et al. 2018). Consistent with the PSI's focus on involving principal supervisors in determining school needs, 71 percent of supervisors in PSI districts agreed that

they were involved in deploying instructional support staff to schools, compared with 41 percent of supervisors in other urban districts.

Figure VIII.4. Principal supervisors in the PSI and other urban districts had similar perceptions of the central office

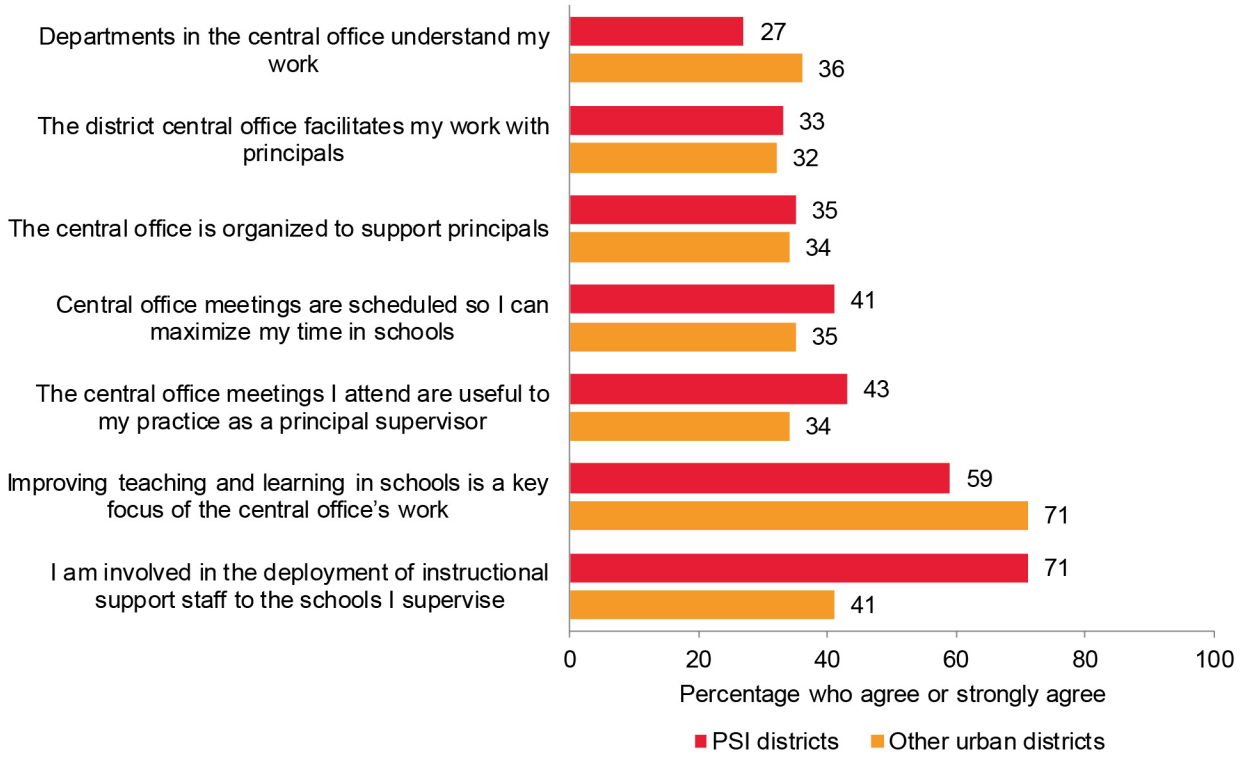


Figure reads: Twenty-seven percent of supervisors in PSI districts agreed or strongly agreed that departments in the central office understand their work.

Source: Wallace Foundation National Survey of Principal Supervisors, 2018 (supervisors in PSI districts, *N* = 50; supervisors in other urban districts, *N* = 293).

Note: The survey question read, "Based on your experiences with this district's central office in the current 2017–2018 school year, how strongly do you agree or disagree with the following statements?" Supervisors rated their agreement on a five-point Likert-type scale. Difference between supervisors in PSI and other urban districts is significant at the 5 percent level for the item "I am involved in the deployment of instructional support staff to the schools I supervise."

D. The work of principal supervisors

Principal supervisors in PSI and other urban districts spent similar amounts of time working with principals. Principal supervisors typically spend their time on a wide variety of activities, including principal coaching and evaluation, district administration, and leading principal professional learning communities. The PSI attempted to further focus each principal supervisor's time on direct interactions with principals and their leadership teams on instructional leadership, rather than operational issues. However, in both PSI and other urban districts, supervisors spent about half their time visiting schools and about 15 percent of their time in network and group meetings with principals, with a total of about 65 percent of their time working and interacting with principals.

Principal supervisors in PSI districts and other urban districts spent similar amounts of time on instructional leadership. In a typical week, principal supervisors in both PSI and other urban districts spent about half of their time with principals focusing on instructional leadership (52 percent in PSI districts and 48 percent in other urban districts) and the other half focusing on other topics.

Principal supervisors in PSI and other urban districts used similar instructional leadership practices. About half of the supervisors in both sets of districts reported that they usually or always used a protocol when discussing data with principals (49 percent in PSI districts and 54 percent in other urban districts) (Figure VIII.5). Similarly, supervisors in both sets of districts were equally likely to report working with principals to assess teachers' effectiveness (69 percent in both sets of districts). However, supervisors in PSI districts were significantly less likely to report that they usually or always provided principals with actionable feedback than those in other urban districts (76 versus 94 percent). This counterintuitive finding could reflect a heightened understanding of how to provide actionable feedback among supervisors in PSI districts, given the PSI's emphasis on this practice. It is possible that supervisors in other urban districts who were not trained in providing actionable feedback might have misunderstood what it entailed and inaccurately believed they provided it frequently.

Figure VIII.5. Principal supervisors in PSI districts and other urban districts engaged in similar instructional leadership practices with principals

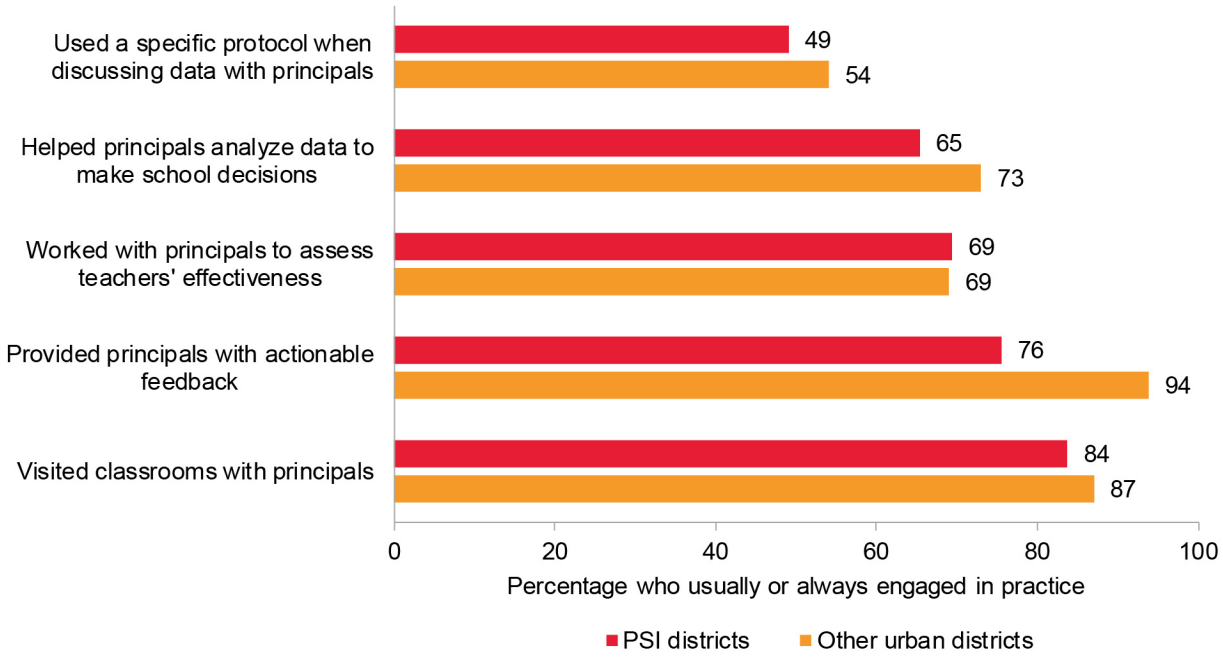


Figure reads: Forty-nine percent of supervisors in PSI districts reported that they usually or always used a specific protocol when discussing data with principals.

Source: Wallace Foundation National Survey of Principal Supervisors, 2018 (supervisors in PSI districts, *N* = 50; supervisors in other urban districts, *N* = 293).

Note: The survey question read, "Thinking about all the time you spent working with principals over the past three months, how often would you say each of the following were true?" Supervisors rated their frequency of practices on a five-point frequency scale. Difference between supervisors in PSI and other urban districts is significant at the 5 percent level for "I provided principals with actionable feedback."

Principal supervisors in PSI districts engaged less frequently in hiring and operational activities than those in other urban districts. Only 10 percent of supervisors in PSI districts reported that they usually or always supported principals with hiring teachers or other school staff, compared with 22 percent of supervisors in other urban districts. Similarly, 20 percent of supervisors in PSI districts indicated that they usually or always helped principals with facilities or other operational issues, compared with 45 percent of supervisors in other urban districts (Figure VIII.6). Some supervisors in PSI districts continued to support principals with annual budgeting, particularly new principals (see Chapter IV, "A spotlight on supervisors' differentiation of support for principals").

Figure VIII.6. Principal supervisors in PSI districts were less likely to work with principals on hiring and operational issues than those in other urban districts

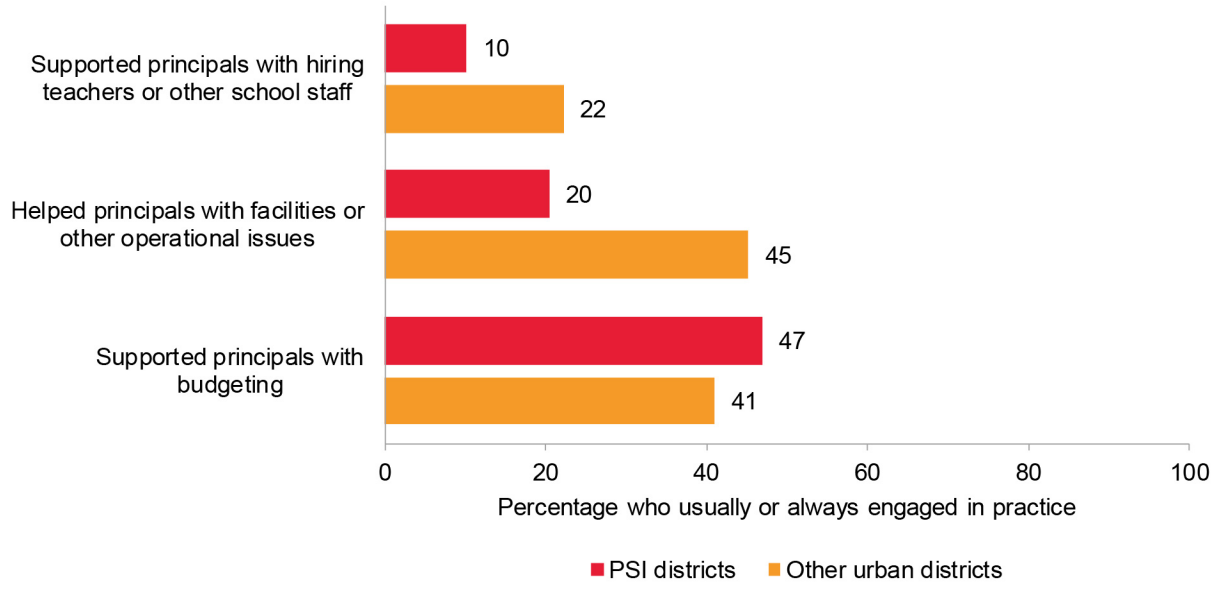


Figure reads: Ten percent of supervisors in PSI districts reported that they usually or always supported principals with hiring teachers or other school staff.

Source: Wallace Foundation National Survey of Principal Supervisors, 2018 (supervisors in PSI districts, *N* = 50; supervisors in other urban districts, *N* = 293).

Note: The survey question read, "Thinking about all the time you spent working with principals over the past three months, how often would you say each of the following was true?" Supervisors rated their frequency of practices on a five-point frequency scale. Differences between supervisors in PSI and other urban districts are significant at the 5 percent level for "I supported principals with hiring teachers or other school staff" and "I helped principals with facilities or other operational issues."

E. Districts' evaluation of principal supervisors and principals

Principal supervisors in PSI districts had more favorable views of their districts' principal evaluation systems than those in other urban districts. The PSI aimed to improve the performance evaluation process for both principals and supervisors as a way to align principal supervision to district standards and specific areas of growth for each principal. About three-quarters (76 percent) of supervisors in PSI districts agreed that the principal evaluation system aligned with the work supervisors and principals did together, compared with 63 percent

of supervisors in other urban districts (Figure VIII.7). Supervisors in PSI districts were also more likely to report that their district's principal evaluation system provided actionable feedback for principals (67 versus 57 percent, although these differences are also not statistically significant at the 5 percent level).

Figure VIII.7. Principal supervisors in PSI districts had more positive views of their districts' principal evaluation systems than those in other urban districts

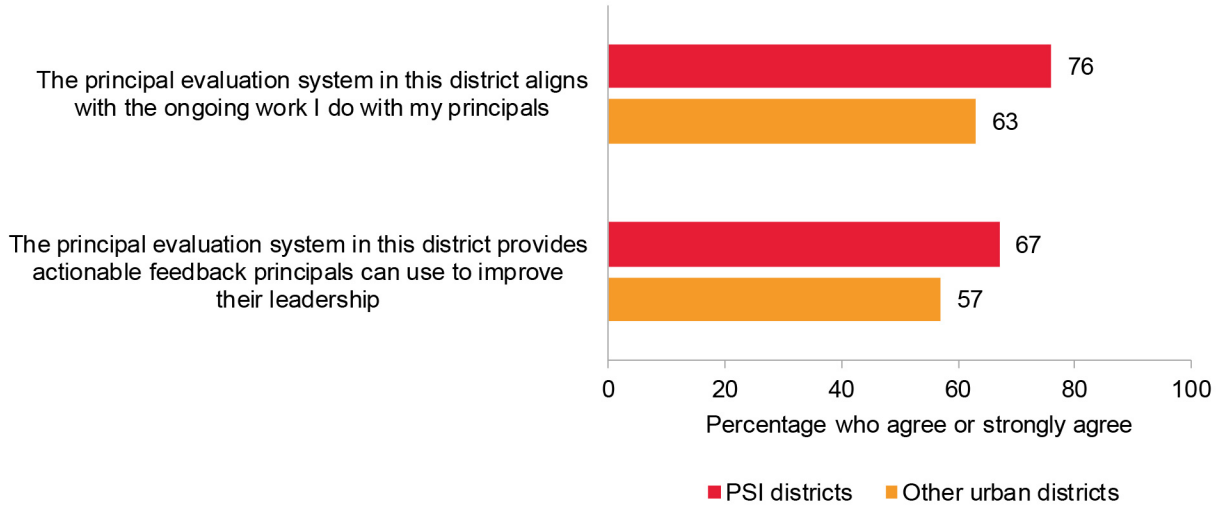


Figure reads: Seventy-six percent of supervisors in PSI districts agreed with the statement "The principal evaluation system in this district aligns with the ongoing work I do with my principals."

Source: Wallace Foundation National Survey of Principal Supervisors, 2018 (supervisors in PSI districts, *N* = 50; supervisors in other urban districts, *N* = 293).

Note: The survey question read, "Based on your experiences in the current 2017–2018 school year, how strongly do you agree or disagree with the following statements about the principal evaluation system in this district?" Supervisors rated their agreement on a five-point Likert-type scale. Differences between supervisors in PSI and other urban districts are not statistically significant.

Principal supervisors in PSI districts also expressed more positive views of their districts' supervisor evaluation systems than did supervisors in other urban districts. Supervisors in PSI districts were more likely than those in other urban districts to agree that their district's principal supervisor evaluation system aligned with their work (55 versus 39 percent), provided actionable feedback (65 versus 41 percent), and was clear (65 versus 53 percent) (Figure VIII.8). They were also more likely to report that their district's principal supervisor evaluation system incorporated principal feedback (57 versus 21 percent).

Figure VIII.8. Principal supervisors in PSI districts had more positive perceptions of their districts' principal supervisor evaluation systems than those in other urban districts

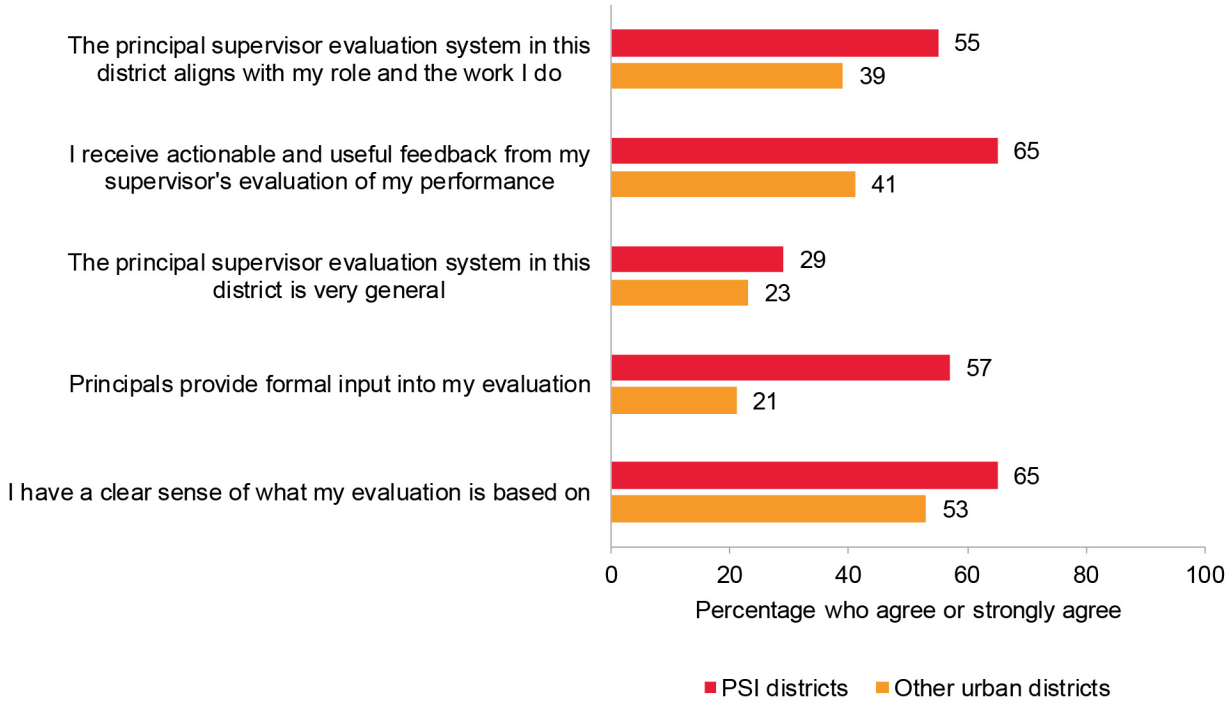


Figure reads: Fifty-five percent of supervisors in PSI districts agreed or strongly agreed with the statement "The principal supervisor evaluation system in this district aligns with my role and the work I do."

Source: Wallace Foundation National Survey of Principal Supervisors, 2018 (supervisors in PSI districts, *N* = 50; supervisors in other urban districts, *N* = 293).

Note: The survey question read, "Based on your experiences in the current 2017–2018 school year, how strongly do you agree or disagree with the following statements about the principal supervisor evaluation system in this district?" Supervisors rated their agreement on a five-point Likert-type scale. Differences between supervisors in PSI and other urban districts are significant at the 5 percent level for "The principal supervisor evaluation system in this district aligns with my role and the work I do," "I received actionable and useful feedback from my supervisors' evaluations of my performance," and "Principals provide formal input into my evaluation."

F. Summary

A comparison of principal supervision in the PSI districts and other urban districts in the final year of the PSI shows several key differences, along with some important similarities. Compared with supervisors in other urban districts, supervisors in the PSI districts had lower spans of control, received more training and mentoring, spent less time on operational issues, and had more favorable views of their districts' principal and supervisor evaluation systems. These differences suggest that the PSI led participating districts to make greater changes to these aspects of principal supervision than the general changes occurring in other urban districts nationwide. However, principal supervisors in PSI and other urban districts spent their time in similar ways and implemented similar instructional leadership practices. These similarities in principal supervisor practices in PSI and other urban districts might have been driven by

principal supervision standards and other national and local efforts to shift the focus of the principal supervisor role in ways that mirrored some of the changes promoted by the PSI.

The similar practices used by principal supervisors in PSI districts and other urban districts could help explain why the study found that the PSI did not improve teachers' perceptions of principals' performance. The study measured the PSI's effects by comparing teachers' perceptions of principals' performance across PSI districts and a set of similar non-PSI districts. If principal supervisors in these other districts engaged in similar activities and practices with their principals, this might explain why the study failed to find any differences in teachers' perceptions of principals' performance, despite the many other changes that the PSI districts made to support supervisors in their work with principals.

IX. CONCLUSIONS AND LESSONS LEARNED

The PSI was a district-level initiative aimed at changing the role of principal supervisors so that they could better support and develop effective principals. The initiative defined five core components that were intended to create the conditions necessary for districts and supervisors to support and develop principals. However, each district adapted these components to their unique contexts and implemented them to best fit their goals.

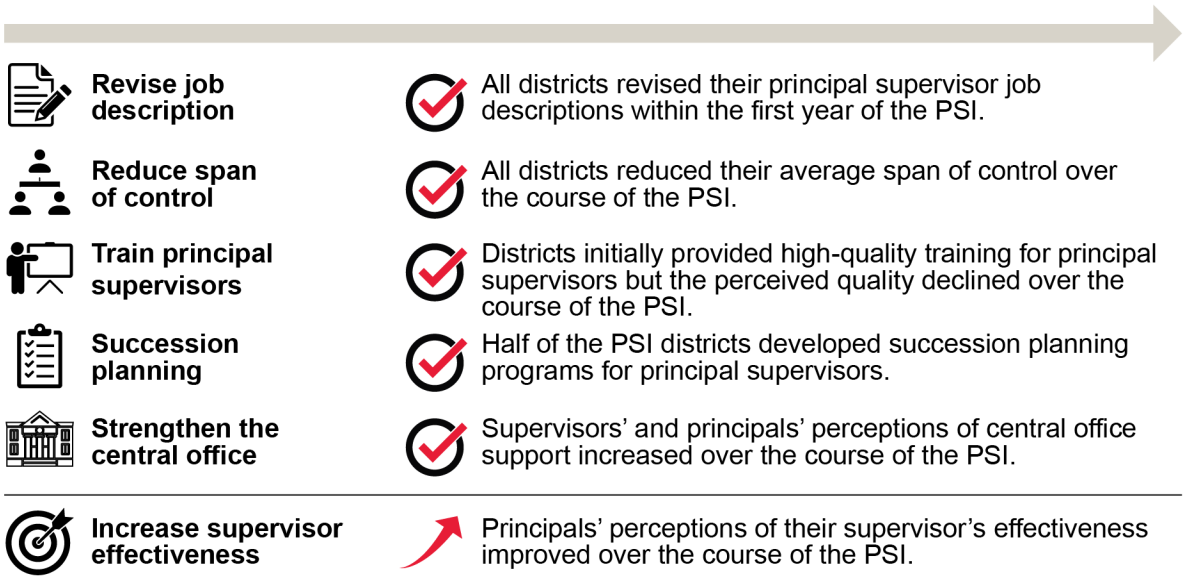
The PSI brought substantial changes to all six districts. Although its ultimate goal was to improve principals' effectiveness as instructional leaders, the PSI did not touch principals directly, but was implemented at the district level. Consequently, some of the most profound changes we observed in the PSI districts occurred within the central office. As they changed the role of the principal supervisor, districts carefully considered the health and stability of their central office organization and made changes to improve the central office's ability to support school improvement as a whole.

The initiative was successful in changing the principal supervisor role to one that focused on providing support and guidance to schools and principals. Supervisors learned specific approaches to supporting principals, such as coaching, and developed a professional community to support their own development. In the revised role, principal supervisors also helped the central office better understand the needs of schools.

The PSI did not improve teachers' perceptions of principals' performance. It is possible the PSI influenced other outcomes not examined by the study, such as principals' ability to hire or retain more effective teachers, or other important outcomes such as principal retention or satisfaction. In addition, the timeframe for the study may have been too short to detect changes in teachers' perceptions of principals' performance resulting from the PSI.

A. Summary of PSI experiences and effects

Districts made substantial changes as a result of the PSI. The PSI's five core components were designed to revise the principal supervisor role and increase supervisor and principal effectiveness. The PSI led to district-level changes in central office culture, structures, and support for schools. Each district faced different challenges and opportunities at the start of the PSI. However, by the end of the initiative, the districts, on average, had made progress along each of the components.



- **Revising the job description.** All districts revised their principal supervisor job description early in the initiative. The job description served as the basis for communicating the change in the principal supervisor role to staff and community members and helped drive the district's vision for principal supervision. The job description served as a guide for districts over the course of the PSI, although none of the districts formally revised it during the initiative.
- **Reducing span of control.** PSI districts reduced their average span of control from 17 to 13 principals and reduced the number of outside responsibilities that could have distracted supervisors from their work with principals. Supervisors with smaller spans of control spent more time meeting with their principals and increased their emphasis on instructional leadership practices during school visits. All supervisors spent the majority of their time working with principals one-on-one or in group meetings.
- **Providing dedicated training for principal supervisors.** In the early years of the initiative, the PSI districts worked hard to provide dedicated training for principal supervisors. As the districts revised the principal supervisor role to focus on improving principal instructional leadership, district leaders learned that they needed to simultaneously define key concepts related to school support. Districts worked to strengthen understanding across the district of high-quality instruction and instructional leadership. They worked with external technical assistance providers to deliver common tools and training to principal supervisors and other central office staff. These trainings were intended to facilitate consistency of practice across supervisors. Districts also aimed to provide a standard of support to principals and schools that was aligned with the district's vision of high-quality instruction, instructional leadership, and school support. In the later years of the initiative, dedicated training waned in some districts in both quantity and quality. In these districts, the focus of meeting times for professional learning and continuous skill development was often overtaken by administrative items and information sharing, leaving new supervisors to learn the complexities of the role on their own.

- **Succession planning.** Half the PSI districts developed specific programs for succession planning for principal supervisors, and other districts without programs became more systematic and comprehensive about their approaches to screening and hiring new principal supervisors. However, districts with aspiring principal supervisor programs faced a challenge in trying to balance the small number of supervisor vacancies each year with a group of aspiring supervisors who had completed a program but yet faced few to no openings in the district. Despite this challenge, over the course of the initiative more supervisors became aware of open principal supervisor positions because of aspiring supervisor programs and opportunities.
- **Strengthening the central office.** Changes to the supervisor role also necessitated both structural and cultural adjustments within the central office. Districts transformed structures and roles to better align with and facilitate principal supervisors' work. They developed new communication systems and approaches. Districts continued to bolster structures they created early on, such as cross-departmental liaisons and central office support teams, to orient the central office toward supporting principals in schools. Supervisors increasingly collaborated with other district departments to plan and coordinate principal professional development and support. Supervisors' and principals' perceptions of central office quality increased steadily over the course of the PSI.

Principal supervisors changed their practices, and practices became more consistent within some districts. Through their shared training and work together, supervisors began to develop a shared professional identity beyond middle management and a universal set of norms and skills to guide their practice. Supervisors' practices included implementing specific coaching models, using protocols for school walk-throughs, and providing feedback to principals. Supervisors worked with principals to help them develop effective teachers through classroom observations and teacher feedback. Some districts strived to improve consistency by sharing common goals for supervisors, standardizing the focus of school visits, adopting common tools to guide supervisors' interactions with principals, working to calibrate evaluation ratings, and setting expectations for the amount of support supervisors should provide to principals and schools.

Principals' perceptions of their work with their supervisor and their supervisor's effectiveness improved. Principals noticed changes in their work with their supervisor over the course of the PSI. This was reflected in their ratings of their supervisor's effectiveness. It was also reflected in their reports of how often their supervisor implemented practices encouraged by the PSI, such as coaching, feedback, and data use. Principals recounted, over the course of the initiative, how their relationships with their supervisors had improved because supervisors better understood their contexts and specific needs; deeper relationships of trust and respect developed. These deeper relationships helped improve the quality and usefulness of the evaluation feedback principals received from their supervisors. Some principals noted that increased support from supervisors expanded their capacity to address instructional matters in their buildings. Some also noted a shift in their own understanding of their roles as instructional leaders.

Despite widespread progress, districts and supervisors faced challenges as they implemented the new supervisor role. The quality of PSI implementation varied both across the six districts and within each district. Many central office level changes, such as increasing

cross-departmental communication and responsiveness to school needs, challenged long-standing organizational culture and context within the district. Some departments did not adjust their operations to accommodate supervisor work, in some cases continuing to make demands on supervisors' time that took them away from principal support. Additionally, supervisors sometimes found the new role itself to be demanding compared with the previous, compliance-oriented role.

The PSI did not improve teachers' perceptions of principals' performance. We found that teachers' perceptions of principals' performance in the PSI districts remained similar to that in a set of similar comparison schools in districts that did not implement the PSI. However, the PSI's effects on teachers' perceptions of principals' performance varied across schools. We found that principals' perceptions of their supervisors' effectiveness and their supervisors' use of practices promoted by the PSI were positively related to the PSI's effects, although this relationship was small. This suggests that the supervisor practices emphasized by the PSI could have some capacity to improve teachers' perceptions of principals' performance. More broadly, the widespread changes districts made during the PSI helped principals feel better supported by the central office and their supervisors, even if the PSI did not lead to direct effects on teachers' perceptions of their performance.

The principal supervisor role in the PSI districts differed from that in other urban districts in several key ways, but there were also some important similarities in supervisors' work with principals. Compared with supervisors in other urban districts, supervisors in the PSI districts had lower spans of control, received more training and mentoring, spent less time on operational issues, and had more favorable views of their districts' principal and supervisor evaluation systems. These differences suggest that the PSI led participating districts to make greater changes to these aspects of principal supervision than the general changes occurring in other urban districts nationwide during the same time frame. However, principal supervisors in PSI and other urban districts spent similar amounts of time working with principals and used similar instructional leadership practices in this work. Principal supervisor standards and other national and local efforts to shift the focus of the principal supervisor role in ways that mirrored some of the changes promoted by the PSI might have driven these similarities. They might also help explain why the study found that the PSI did not improve teachers' perceptions of principals' performance, relative to a similar set of non-PSI districts.

B. Looking ahead: Sustaining the changes in the PSI districts

Some of the six PSI districts will continue to focus on supporting and developing the principal supervisor role to better support principals, while others will undoubtedly move away from this focus as a major district initiative. Our findings suggest that the following actions may be important for sustainability of PSI-driven changes.

Embedding the principal supervisor role within the broader structures and work of the central office. Some districts connected central office staff and principal supervisors through structures such as cross-departmental and cross-functional teams, as well as intentional trainings and meetings that met the needs of both groups. These structures facilitated ongoing communication and relationship building. Principal supervisors in many districts took on

leadership of districtwide school improvement efforts. These new approaches to central office and supervisor interactions can continue to build sustainability of the PSI.

Communicating the importance of high-quality principal support to stakeholders. Some districts were very intentional about explaining the purpose and benefits of the PSI to department chiefs and school board members. Stakeholders who understood how the PSI aligned to district goals were more supportive of efforts to reallocate roles and resources around principal supervision. Ongoing communication and explanations of the work can help sustain the achievements of the PSI.

Obtaining financial resources to support PSI changes. Districts obtained financial allocations to fully fund the added supervisor positions and transition from external to in-house training for principal supervisors. Those districts that have funded the PSI for years to come will be able to sustain the initiative's momentum.

Developing an understanding among senior district leaders that the principal supervisor role is not static. Several districts planned to revisit and refresh the principal supervisor job description and responsibilities as needed to match evolving district goals for school support, thus ensuring the long-term relevance of the role. This understanding is central to the ongoing goals of the PSI.

Championing of the PSI and its vision by the superintendent after the initiative ends. Districts with superintendents who remained engaged with the PSI saw stronger implementation and were better set up for sustainability than those that experienced leadership changes and turnover. Superintendents moved the work of the initiative along by consistently monitoring progress, engaging in dialogue with principal supervisors and their departments, and communicating the importance of the change to district stakeholders. As the initiative drew to a close, these superintendents could clearly articulate how they planned to continue and scale up the work of the PSI within their districts. Superintendents spoke of ensuring sustainability of the PSI by (1) safeguarding reduced span of control by securing necessary funding, (2) aligning principal supervision with districtwide priorities and initiatives, (3) continuing to signal support of the PSI work to stakeholders such as school board members, and (4) articulating a commitment to ongoing central office reform to ensure departments were focused on providing support to principals and schools.

C. Lessons learned for revising the principal supervisor role

The experiences of the PSI provide lessons for other districts to consider when implementing district-level efforts to revise the principal supervisor role.

> **Make sure the role is positioned in a way that it can actually have access and authority to resources for schools right off the bat, so you're not fighting for it or vying over territory issues. And then I would say make it a declared priority from the superintendent of the value of that position in the organization.**
>
> ***—District leader (2018)***

Obtain buy-in and build awareness across all stakeholders. District leaders repeatedly referenced the importance of building buy-in and awareness across all stakeholders in the district, including board members, central office staff, and principals, to ensure the success of the PSI. Several central office staff spoke of the importance of involving top-level leaders of other central office departments as well as school board members to support and provide resources for the work of the PSI.

> **I would say that we have to get a little bit better at identifying specific needs … because the instructional piece I feel like we get a lot of, but maybe [we need] supports around school culture and getting more parent involvement and those types of things that aren't necessarily instruction.**
>
> ***—Principal (2018)***

Balance supervisors' focus on instructional leadership with the flexibility to meet principal needs. While a main thrust of the PSI was to focus supervisors on supporting principals' instructional leadership, this was not meant to be supervisors' only focus. Supervisors, principals, and central office staff described the importance of building flexibility into the role to meet the wide-ranging needs of principals and their schools. For instance, principal supervisors described the need to spend far more time on logistical and operational issues as well as instructional leadership with new principals. Districts worked to balance the need for supervisors to differentiate their work across principals. A principal described the importance of needs-based differentiation.

> **It's not just creating a fiefdom in a charismatic former principal who now wants to lead principals, [it's] helping those who supervise principals to see themselves as part of the [district as a] whole.**
>
> ***—District leader (2018)***

Invest in selecting and training the best candidates for the principal supervisor position. District officials and supervisors repeatedly emphasized the importance of selecting the right individuals for the principal supervisor role beyond just choosing "superstar principals." Many agreed that prior principal experience was an asset to success in the role. However, leaders in all districts noted that the supervisor role was dramatically different from the principalship and required a diverse set of skills. In addition to identifying high-performing principals as potential supervisors, districts also looked for teaching-oriented individuals who were reflective, good listeners, skilled at data use, and oriented to the district's mission.

Consider the trade-offs in different strategies for assigning principals to supervisors. As districts reduced principal supervisors' span of control, they employed different strategies to assign supervisors to schools. Typically, districts assigned supervisors to schools based on one

primary factor, such as grade levels or geography, but they also considered secondary factors, such as school performance. Some districts experimented with placing their lowest-performing schools together. By the PSI's final year, many principals and supervisors had formed opinions about the strengths and weaknesses of particular approaches but had not reached a consensus.

From the principal supervisors' perspective, geographic organization of schools was beneficial because it reduced the amount of time they spent traveling between schools. Some principals noted that they enjoyed being in a network with other schools in the same geographic area because they tended to have other features in common, such as student demographics and community. Similarly, principals in schools grouped according to specialty or grade level appreciated the opportunity to work with principals who had similar specializations. A principal in Broward explained: *"*I'd rather spend [principal meeting] time on similar issues with schools like mine." Principals highly valued supervisors who had prior experience in a school at the same grade level as the school the principal was leading.

At the same time, some principals in networks of similar schools expressed the desire to have more opportunities to learn from principals who could offer different perspectives. This desire was often voiced by principals grouped into "turnaround" or "priority" networks, who wanted to learn best practices from principals in higher-performing schools. Some central office personnel noted that although these turnaround groups were conceived as ways to concentrate resources in schools, they were one of the most challenging kinds of networks for supervisors.

Networks with diverse groups of schools also had strengths and weaknesses. Supervisors were sometimes challenged to tailor their supports to a wide range of school and principal needs. Higher-performing principals in diverse networks often complained that network-based activities were not helpful because they were directed to the needs of lower-performing schools.

Consider the stability of district context and leadership. District stability was important for positive PSI experiences. In some districts, supervisors and central office staff explicitly noted that their ability to address challenging aspects of the PSI was clouded by the context they faced. This context included superintendent changes; turnover of top-level central office leaders; deep resource and financial constraints; and a lack of overall stability, including unfilled positions, diminishing enrollments, community trauma, and bureaucratic tensions throughout the school system. These issues also often resulted in the PSI losing its champion, and made it difficult for district personnel to understand the new principal supervisor role. Principal supervisors in districts facing less stability felt that there was not a consistent vision or set of expectations for the new role. They became bogged down in operational and maintenance issues, or providing basic supports to schools, because other departments were stretched too thin.

Change structures and transform values, beliefs, and behaviors. Changing the supervisor role required changes to the structures in the central office. However, it also required adaptive change—or changing values, beliefs, and behaviors in the district. District staff had to change how they envisioned the work of the supervisor, how they interacted with other personnel within the district, and how they understood their own roles in the district. While structural changes can be made relatively quickly, adaptive changes take longer and require ongoing attention and reinforcement. Breaking down divisions and creating mechanisms to facilitate communication require cultural shifts so that individuals can work across departments. Without

this adaptive change, cross-departmental work is often territorial and becomes threatening to staff. Furthermore, as supervisors work to ensure that the central office supports their principals, they challenge long-held beliefs that schools should be accountable to the central office. Under the revised supervisor role, while schools had to be accountable to the central office, the central office had to be accountable to schools as well.

D. Questions for further research

Future research should continue to explore the effectiveness of efforts to leverage the principal supervisor role to support and develop principals as instructional leaders. Many of the changes to the supervisor role promoted by the PSI, including reductions in span of control and increases in support and training for principal supervisors, have been adopted to varying degrees by urban school districts throughout country (Cochran et al. 2020). These broader national trends, along with varying approaches to principal supervision across a diverse set of districts, will provide additional opportunities to explore how best to leverage the principal supervisor role to improve principal's performance. For example, research might examine district-level efforts to revise the principal supervisor role in conjunction with implementation of a principal pipeline to develop, select, and support principals. Future research could also examine the effectiveness of such efforts on a broader set of outcomes beyond teachers' perceptions of principals' performance, including principal retention, teacher retention and satisfaction, and student achievement. Finally, future research might continue to explore the relationship between teachers', principals', and supervisors' perceptions of principals' performance.

REFERENCES

Anderson, E., and M.D. Young. "A Research-Based Framework for District Effectiveness." *UCEA Review,* fall, 2018a, pp. 1–11.

Anderson, E., and M.D. Young. "If They Knew Then What We Know Now, Why Haven't Things Changed? An Examination of District Effectiveness Research." *Frontiers in Education*, vol. 3, 2018b, pp. 1–20.

Boyd, D., P. Grossman, M. Ing, H. Lankford, S. Loeb, and J. Wyckoff, J. "The Influence of School Administrators on Teacher Retention Decisions." *American Educational Research Journal*, vol. 48, no. 2, 2011, pp. 303–333.

Burch, P., and J. Spillane. "Leading from the Middle: Mid-Level District Staff and Instructional Improvement." Chicago, IL: Cross City Campaign for Urban School Reform, 2004.

Busso, M., J. DiNardo, and J. McCrary. "New Evidence on the Finite Sample Properties of Propensity Score Reweighting And Matching Estimators." *Review of Economics and Statistics*, vol. 96, no. 5, 2014, pp. 885–897.

Cameron, A.C., and D.L. Miller. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources*, vol. 50, no. 2, 2015, pp. 317–372.

Cameron, A.C., J.B. Gelbach, and D.L. Miller. "Bootstrap-Based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics*, vol. 90, no. 3, 2008, pp. 414–427.

Cicchetti, D.V. "Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology." *Psychological Assessment,* vol. 6, no. 4, 1994, p. 284.

Cochran, A., R. Hart, M. Casserly. "Trends in Principal Supervisor Leadership and Support: Results from Two Surveys of Principal Supervisors in America's Great City Schools." Washington, DC: Council of the Great City Schools, 2020. Available at https://www.wallacefoundation.org/knowledge-center/Pages/Trends-in-Principal-Supervisor-Leadership-and-Support.aspx.

Cohen, J. "A Coefficient of Agreement for Nominal Scales*." Educational and Psychological Measurement,* vol. 20, 1960, pp. 37–46.

Corbin, J.M., and A.L. Strauss. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory.* Thousand Oaks, CA: Sage, 2008.

Corcoran, A., M. Casserly, R. Price-Baugh, D. Walston, R. Hall, and C. Simon. "Rethinking Leadership: The Changing Role of Principal Supervisors." Washington, DC: Council of Great City Schools, 2013.

Council of Chief State School Officers. "Model Principal Supervisor Professional Standards 2015." Washington, DC: CCSSO, 2015.

Covay Minor, E., A.C. Porter, J. Murphy, E.B. Goldring, X. Cravens, and S.N. Elliot. "A Known Group Analysis Validity Study of the Vanderbilt Assessment for Leadership in Education in U.S. Elementary and Secondary Schools." *Educational Assessment, Evaluation and Accountability,* vol. 26, no. 1, 2014, pp. 29–48.

DeVellis, R.F. *Scale Development: Theory and Applications* (vol. 26). Los Angeles: Sage, 2012.

Fleiss, J.L. "Measuring Nominal Scale Agreement Among Many Raters." *Psychological Bulletin,* vol. 76, no. 5, 1971, p. 378.

Frölich, M. "Finite-Sample Properties of Propensity-Score Matching and Weighting Estimators." *Review of Economics and Statistics,* vol. 86, no. 1, 2004, pp. 77–90.

Gates, S.M., M.D. Baird, C.J. Doss, L.S. Hamilton, I.M. Opper, B.K. Master, A.P. Tuma, M. Vuollo, and M.A. Zaber. "Preparing School Leaders for Success**:** Evaluation of New Leaders' Aspiring Principals Program, 2012–2017." Santa Monica, CA: RAND Corporation, 2019a. Available at https://www.rand.org/pubs/research_reports/RR2812.html. Accessed June 17, 2019.

Gates, S.M., M.D. Baird, B.K. Master, and E. Chavez-Herrerias. "Principal Pipelines: A Feasible, Affordable, and Effective Way for Districts to Improve Schools." New York: The Wallace Foundation, 2019b. Available at https://www.wallacefoundation.org/knowledge-center/Documents/Principal-Pipelines-A-Feasible-Affordable-and-Effective-Way-for-Districts-to-Improve-Schools.pdf. Accessed June 17, 2019.

Goff, P., J. Guthrie, E. Goldring, and L. Bickman. "Changing Principals' Leadership Through Feedback and Coaching." *Journal of Educational Administration,* vol. 52, no. 5, 2014, pp. 682–704.

Goldring, E., X. Cravens, J. Murphy, A. Porter, and S. Elliott. "The Convergent and Divergent Validity of the Vanderbilt Assessment of Leadership in Education (VAL-ED): Instructional Leadership and Emotional Intelligence." *Journal of Educational Administration,* vol. 53, no. 2, 2018 pp. 177–196.

Goldring, E.B., J.A. Grissom, M. Rubin, L.K. Rogers, M. Neel, and M.A. Clark. "A New Role Emerges for Principal Supervisors: Evidence from Six Districts in the Principal Supervisor Initiative." New York: Wallace Foundation, 2018. Available at https://www.wallacefoundation.org/knowledge-center/Documents/A-New-Role-Emerges-for-Principal-Supervisors.pdf. Accessed September 15, 2019.

Grissom, J.A., and S. Loeb. "Triangulating Principal Effectiveness: How Perspectives of Parents, Teachers, and Assistant Principals Identify the Central Importance of Managerial Skills." *American Educational Research Journal,* vol. 48, no. 5, 2011, pp. 1091–1123.

Grissom, J.A., D. Kalogrides, and S. Loeb. "Using Student Test Scores to Measure Principal Performance." *Educational Evaluation and Policy Analysis,* vol. 37, no. 1, 2015, pp. 3–28.

Hall, P., J. Racine, and Q. Li. "Cross-Validation and the Estimation of Conditional Probability Densities." *Journal of the American Statistical Association*, vol. 99, no. 468, 2004, pp. 1015–1026.

Hallinger, P., L. Bickman, and K. Davis. "School Context, Principal Leadership, and Student Reading Achievement." *The Elementary School Journal,* vol. 96, no. 5, 1996, pp. 527–549.

Honig, M.I., M. A. Copland, L. Rainey, J.A. Lorton, and M. Newton. "Central Office Transformation for District-Wide Teaching and Learning Improvement." Seattle, WA: Center for the Study of Teaching and Policy, University of Washington, 2010.

Huber, M., M. Lechner, and C. Wunsch. "The Performance of Estimators Based on the Propensity Score." *Journal of Econometrics*, vol. 175, no. 1, 2013, pp. 1–21.

Imai, K., G. King, and E.A. Stuart. "Misunderstandings Among Experimentalists and Observationalists in Causal Inference." *Journal of the Royal Statistical Society,* Series A, vol. 171, 2008, pp. 481–502.

Jacob, R., R. Goddard, M. Kim, R. Miller, and Y. Goddard. "Exploring the Causal Impact of the McREL Balanced Leadership Program on Leadership, Principal Efficacy, Instructional Climate, Educator Turnover, and Student Achievement." *Educational Evaluation and Policy Analysis,* vol. 52, 2014, pp. 187–220.

Landis, J. Richard, and Gary G. Koch. "The Measurement of Observer Agreement for Categorical Data." *Biometrics*, vol. 33, no. 1, 1977, pp. 159–174

LeCompte, M.D., and J.J. Schensul. *Analyzing and Interpreting Ethnographic Data*. Lanham, MD: AltaMira Press, 1999.

Marsh, J.A., K.A. Kerr, G.S. Ikemoto, H. Darilek, M. Suttorp, R.W. Zimmer, and H. Barney. "The Role of Districts in Fostering Instructional Improvement Lessons from Three Urban Districts Partnered with the Institute for Learning." Santa Monica, CA: RAND Corporation, 2005. Available at http://eric.ed.gov/?id=ED486658. Accessed June 17, 2019.

Neumerski, C.M., J.A. Grissom, E. Goldring, T. Drake, M. Rubin, M. Cannata, and P. Schuermann. "Restructuring Instructional Leadership: How Multiple-Measures Teacher Evaluation Systems Are Redefining the Role of the School Principal." *The Elementary School Journal*, vol. 119, no. 2, October 2018.

Porter, A.C., J. Murphy, E. Goldring, S.N. Elliot, M.S. Polikoff, and H. May. *Vanderbilt Assessment of Leadership in Education: Technical Manual, Version 1.0.* Nashville, TN: Learning Sciences Institute, Vanderbilt University, November 2008. Available at https://www.wallacefoundation.org/knowledge-center/Documents/Vanderbilt-Assessment-of-Leadership-in-Education-Technical-Manual-1.pdf. Accessed May 8, 2019.

Porter, A.C., M.S. Polikoff, E. Goldring, J. Murphy, S.N. Elliott, and H. May. "Investigating the Validity and Reliability of the Vanderbilt Assessment of Leadership in Education." *Elementary School Journal,* vol. 111, no. 2, 2010a, pp. 282–313.

Porter, A.C., M.S. Polikoff, E. Goldring, J. Murphy, S.N. Elliott, and H. May. "Developing a Psychometrically Sound Assessment of School Leadership: The VAL-ED as a Case Study." *Educational Administrative Quarterly,* vol. 46, no. 2, 2010b, pp. 135–173.

Porter, A.C., R. Yang, J. Caines, E. Goldring, J. Murphy, and S. Elliott. "Using VALED to Predict Value-added to Student Achievement." Presented at the American Educational Research Association Annual Meeting, Philadelphia, PA, April 2014.

Rubin, Donald B. "Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation." *Health Services & Outcomes Research Methodology,* vol. 2, nos. 3–4, 2001, pp. 169–188.

Sebastian, J., and E. Allensworth. "The Influence of Principal Leadership on Classroom Instruction and Student Learning: A Study of Mediated Pathways to Learning." *Educational Administration Quarterly,* vol. 48, no. 4, 2012, pp. 626–663.

Silverman, B.W. *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall, 1986.

Smith, J., and P. Todd. "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics,* vol. 125, nos. 1–2, 2005, pp. 305–353.

Steiner, P.M., T.D. Cook, and W.R. Shadish. "On the Importance of Reliable Covariate Measurement in Selection Bias Adjustments Using Propensity Scores." *Journal of Educational and Behavioral Statistics*, vol. 36, no. 2, 2011, pp. 213–236.

Stuart, E.A. "Matching Methods for Causal Inference: A Review and a Look Forward." *Statistical Science,* vol. 25, no. 1, 2010, pp. 1–21.

APPENDIX A.

ADDITIONAL DETAILS ON METHODS

This appendix provides additional details on the study methods described in Chapter II for the analyses of (1) district, supervisor, and principal experiences with the PSI; (2) the PSI's effects on teachers' perceptions of principals' performance; (3) variation in teachers' perceptions of principals' performance and the PSI's effects across districts and supervisors; and (4) factors associated with the PSI's effects.

A. Examining district, supervisor, and principal experiences with the PSI

To study districts', supervisors', and principals' experiences with the PSI, we conducted site visits and collected surveys in the six PSI districts during the second, third, and fourth years of the initiative (the 2015–2016 through 2017–2018 school years). Chapter II provides an overview of these data and how we analyzed them. Here we provide additional information on the analysis of data from site visits and the survey data and analysis.

1. Analysis of data from site visits

The analysis of data from site visits followed a multistep process. The research team developed a coding framework reflecting the components of the PSI and themes that emerged from each of the site visits. Following each site visit, researchers also developed field notes in structured post-visit reflection forms. The development of the coding framework was iterative (Corbin and Strauss 2008; LeCompte and Schensul 1999); members of the research team coded a sample of central office, supervisor, and principal interviews and then revised the coding scheme to address questions and concerns, until the team arrived at a comprehensive framework.

Researchers used the program Dedoose[12] to code and analyze transcribed interviews. To ensure interrater reliability, we used the code application test function in Dedoose to calculate a pooled Cohen's kappa value (Cohen 1960). Although there are a variety of proposed standards for evaluating the pooled Cohen's kappa value, the score range of our coders fell comfortably within accepted bounds (Landis and Koch 1977; Cicchetti 1994; Fleiss 1971). Scores among the five coders for this study ranged from 0.71 to 1.00. After ensuring there was sufficient interrater reliability, team members coded the interview transcripts.

The research team then analyzed the coded data and developed district case studies. They then analyzed the district case studies to identify emergent cross-district themes, patterns, and district-level variation. Collectively, these findings became the basis of the qualitative findings in this report.

[12] Dedoose, version 8.0.35, is a web application for managing, analyzing, and presenting qualitative and mixed method research data (SocioCultural Research Consultants, LLC, www.dedoose.com).

2. Survey administration

We administered online surveys to principals and principal supervisors in each PSI district three times over the course of the initiative:

- Round 1 (second year of PSI): November 2015–February 2016
- Round 2 (third year of PSI): March–June 2017
- Round 3 (fourth year of PSI): April–July 2018

Due to the timing of the evaluation, we were unable to administer surveys in the first year of the PSI; however, the round 1 principal supervisor survey included some questions asking supervisors to recall their experiences from the previous school year.

Central office staff in the study districts worked with the study team to help ensure high response rates. For example, they typically let principals and supervisors know that they would be receiving a confidential survey from the study team and encouraged them to complete it. They also sent reminders over the course of the survey administration period to help boost response rates. In some cases, districts set aside time during a district meeting for principals and supervisors to complete the survey. As a result, response rates were high across all three rounds of the survey, ranging from 92 to 94 percent for the principal survey and 96 to 100 percent for the supervisor survey (see Table II.2).

3. Survey contents

The principal and principal supervisor surveys included questions about respondents' experiences with each component of the PSI, principal supervision, and changes in the central office, as well as their demographic characteristics (Table A.1). Supervisor surveys also included questions about employment background and hiring experience. The questions stayed largely consistent across rounds to facilitate longitudinal analyses. However, in each year, we added or removed items based on what we learned about the PSI from the previous year's data analysis.

In rounds 2 and 3 of the survey, we added items to assess the extent to which supervisors differentiated their support for principals, based on The Wallace Foundation's interest in this topic. Specifically, for each supervisor, we asked about two specific principals in their network—the most and least experienced. In cases where there were multiple principals with the same years of experience among the most and least experienced principals, we also considered school achievement in the selection. Specifically, from the set of the most experienced principals we selected one from a high-achieving school. From the set of the least experienced principals, we selected one from a low-achieving school.[13] We excluded principals who were new to the district that year from the sampling frame.

Many survey questions included a specific timeframe to ensure consistency of responses. For example, questions about the frequency of various principal supervisor practices asked how frequently these practices had occurred "over the past three months," or roughly one semester.

[13] We classified schools as high- or low-achieving based on whether their achievement was above or below the district median, based on publicly available school achievement data from state websites.

Other questions referred to "the current school year," with the exception of supervisor professional development questions, which referred to the current school year *and* the preceding summer. Because we adminstered the round 1 survey after the PSI had already started, we included questions in the round 1 supervisor survey about supervisors' experiences in the prior school year. Only supervisors who indicated that they had worked in the role in the previous year answered those questions. Also, because we administered the round 1 surveys in the fall, we did not include professional development questions from the current school year (2015–2016) in those surveys, as supervisors and principals would not have had much opportunity to experience professional development related to the PSI.

Table A.1. Topics included in principal and principal supervisors surveys, by round of survey

	Principal supervisor				Principal		
	Round 1		Round 2	Round 3	Round 1	Round 2	Round 3
Topic	2014–2015 recall	2015–2016	2016–2017	2017–2018	2015–2016	2016–2017	2017–2018
Background, demographics, and hiring		X	X[a]	X[a]	X	X[a]	X[a]
Supervisor role and practices	X	X	X	X	X	X	X
Professional development	X		X	X			
Span of control	X	X	X	X	X	X	X
Principal supervisor selection and aspiring supervisor program	X	X	X	X			
Central office redesign	X	X	X	X	X	X	X
Supervisor efficacy and effectiveness		X		X	X	X	X
Differentiation			X	X		X	X

Note: Only supervisors who indicated they had worked in the role in the previous year received 2014–2015 recall questions. 2014–2015 was the first year of the initiative for all districts except Baltimore.

[a] Asked only of new respondents.

4. Constructing measures of principal supervisors' emphasis on key practices

For the analysis of supervisors' differentiation of support across principals, we used responses from the principal supervisor survey on the supervisors' work with their most experienced principal in a high-achieving school and their least experienced principal in a low-achieving school. We then created measures of the level of emphasis each supervisor placed on four sets of key practices in their work with these two principals: (1) data use, (2) classroom visits, (3) feedback, and (4) instruction. Table A.2 shows the items included in each measure. To calculate the supervisor's level of emphasis in each category, we summed the supervisor's response to the set of questions in that category for each of the two principals he or she reported on.

Table A.2. Survey items included in the composite measures of supervisor's emphasis on key practices

Survey items in each measure
Thinking about your visits to the principal at X, how much emphasis do you place on the following practices? (0 = no emphasis – 10 = maximum emphasis)
Emphasis on data use (Cronbach's alpha = 0.87)
Reminding the principal to look at their school's data frequently
Helping the principal think through their school's data and what the data say about school needs
Challenging the principal to plan specific actions or next steps for their school based on data
Supporting the principal's ongoing use of data to make continuous adjustments to school practices in pursuit of long-term goals
Emphasis on classroom visits (Cronbach's alpha = 0.79)
Reminding the principal to conduct informal classroom visits
Helping the principal fit informal classroom visits into their schedule
Helping the principal focus on specific "look-fors" or specific instructional indicators when they make informal classroom visits
Working with the principal to translate what they see in classroom visits into specific feedback for teachers
Emphasis on feedback (Cronbach's alpha = 0.79)
Discussing with the principal the importance of providing timely and frequent feedback to teachers (beyond what is required for formal observations)
Modeling or role-playing teacher feedback with the principal
Observing the principal giving feedback to teachers to help the principal make adjustments or refinements to improve the effectiveness of the feedback
Coaching the principal in having difficult or "courageous" conversations with teachers
Emphasis on instruction (Cronbach's alpha = 0.72)
Developing the principal's understanding of effective instructional practices
Guiding the principal in planning schoolwide professional development program that is appropriate for the specific needs of the school
Helping the principal refine and differentiate professional learning opportunities for teachers according to each teacher's instructional improvement needs

5. Constructing measures of PSI implementation, supervisors' practices, and supervisors' effectiveness

To measure the implementation of individual PSI components, as well as supervisor practices and supervisor effectiveness, we constructed scales using data from the principal and principal supervisor surveys. We use these measures both in the descriptive analysis of PSI experiences in Chapter IV and in the analysis of factors correlated with PSI effects in Chapter VII. Tables A.3 through A.5 show the survey items that contributed to each composite measure. To create the scales, we first reverse-coded any negatively worded items and then calculated the average for all items in the scale.

Table A.3. Survey items included in the composite measures of supervisor training

Survey items in each measure
Training quality (principal supervisor survey) (Cronbach's alpha = 0.98)
Thinking about the district-sponsored training or professional development for principal supervisors you attended during the [year] school year and the summer before, to what extent do you agree or disagree with each of the following statements (strongly disagree, disagree, neither agree nor disagree, agree, or strongly agree)?
Provided opportunities to share specific practices with other principal supervisors
Provided opportunities for me to receive feedback on my practice
Addressed real challenges I face in my role
Gave me opportunities to plan my work
Was based on problems of practices I face in my role*
Provided opportunities for self-assessment of my skills
Provided me tools to set goals for my own development
Was part of a sustained, systematic program for my development
Allowed me to model practices I had learned
Specifically enhanced my capacity to develop principals' instructional leadership
Was engaging
Provided me with actionable tools/resources that I can use in my role
Was interactive
Taught me new knowledge and/or skills
Stimulated my interest
Helped build a learning community with my fellow supervisors
Facilitated my overall leadership
Was geared toward implementing district initiatives and programs
Helped me understand district procedures
Extent to which training aligned with the PSI (principal supervisor survey) (Cronbach's alpha = 0.87)
Thinking about the district-sponsored training or professional development for principal supervisors you attended during the [year] school year and the summer before, how much emphasis was placed on the following areas (no emphasis, some emphasis, or a great deal of emphasis)?
Skills for coaching principals
Working effectively one-one-one with principals
Providing actionable and specific feedback to principals
Differentiating support for principals
Conducting difficult conversations
Coaching principals on giving teachers actionable feedback

Table A.4. Survey items included in the composite measures of central office support

Survey items in each measure
Quality of central office support (principal supervisor survey) (Cronbach's alpha = 0.90)
Please indicate how strongly you agree with each of the following statements about the central office in this district (strongly disagree, disagree, neither agree nor disagree, agree, or strongly agree).
The district central office facilitates my work with principals.
Central office meetings are scheduled so I can maximize my time in schools.
The way the central office is organized interferes with my ability to work with principals.
The way the central office is organized interferes with my ability to collaborate with other principal supervisors.
Departments in the central office understand my work.
The central office meetings I attend are useful to my practice as a principal supervisor.
Improving teaching and learning in schools is a key focus of the central office's work.
My principals seek help from me because they do not know who to contact in the central office.
Turnover at the central office interferes with the ability of departments to assist principals and schools.
Principals lose time focusing on teaching and learning because of central office requests.
The central office is organized to support principals.
Quality of central office support (principal survey) (Cronbach's alpha = 0.90)
Please indicate to what extent you agree or disagree with the following statements (strongly disagree, disagree, neither agree nor disagree, agree, or strongly agree).
I know who to contact in the central office when I need assistance.
Departments in the central office efficiently coordinate school-related work.
The central office efficiently provides the services I require for my school.
Improving teaching and learning in schools is a key focus of the central office's work.
Departments in the central office reach out to me to help me with my needs.
Working with the departments in the central office is difficult because departments do not work together.
I seek help from my principal supervisor because the central office doesn't respond to my requests.
I seek help from my principal supervisor because I do not know who to contact in the central office.
Turnover at the central office interferes with the ability of departments to assist principals and schools.
I lose time focusing on teaching and learning because of requests from the central office.
The central office is organized to support me in my role as principal.

Table A.5. Survey items included in the composite measures of supervisor effectiveness

Survey items in each measure
Supervisors' effectiveness as perceived by principals (principal survey) (Cronbach's alpha = 0.98)
How effective would you say your principal supervisor is at each of the following (not at all effective, not very effective, somewhat effective, effective, or very effective)?
Helping me assess my strengths and weaknesses
Holding me accountable for taking specific steps or action
Assisting me with setting goals
Ensuring I am implementing district policies or priorities
Making sure I respond to central office requests
Holding me accountable for my evaluation results
Providing me with actionable feedback
Monitoring my development
Advocating for my needs as a principal with district leaders
Buffering me from outside interference
Connecting me with other central office personnel when needed
Garnering resources for me
Linking me to district or external expertise when needed
Helping me improve my teachers' instruction
Helping me focus my time on instruction and teaching
Improving the quality of feedback I give my teachers
Helping me use and understand my school's data
Creating a professional learning community for me and other principals
Helping me raise student achievement in my school
Encouraging me
Being a cheerleader for me
Helping me improve my work-life balance
Developing a trusting relationship with me
Supporting me during difficult situations
Addressing parent or community member concerns
Assisting me with school budgeting
Assisting me with school facilities issues
Helping me meet the needs of diverse learners
Helping me implement challenging curricula and assessments
Extent to which supervisors implemented the specific practices that were the focus of the PSI (principal survey) (Cronbach's alpha = 0.93)
Over the past three months, thinking about all the time you spent working with your principal supervisor, how often would you say each of the following were true (never, rarely, sometimes, usually, or always)?
My principal supervisor used a specific coaching approach or model with me.
My principal supervisor helped me analyze data to make school decisions.
My principal supervisor used data to set goals for his or her work with me.
My principal supervisor helped me align the school's budget with school priorities.
My principal supervisor worked with me to assess my teachers' effectiveness.
My principal supervisor provided me with actionable feedback.
My principal supervisor used a system for monitoring my growth and change from one visit to the next.
I was informed in advance of principal supervisor visits to school.
My principal supervisor and I jointly decided on goals for visits to school.
My principal supervisor developed a specific agenda in advance of visits to school.
My principal supervisor communicated the goals for our work during the visit.
My principal supervisor documented what we discuss during a school visit.
When my principal supervisor visited my school we worked on whatever I was doing at the time of the visit.

Survey items in each measure
My principal supervisor modeled effective teaching practices.
My principal supervisor modeled effective feedback and coaching.
My principal supervisor role-played practices he or she hoped to see in my school.
My principal supervisor worked directly with teachers in my school.
My principal supervisor worked directly with assistant principals, coaches, or other school leaders.
Quality of supervisors' evaluation of and feedback for principals (principal survey) (Cronbach's alpha = 0.94)
Thinking about the feedback you received on your performance evaluation from your principal supervisor, please indicate to what extent you agree or disagree with the following statements (strongly disagree, disagree, neither agree nor disagree, agree, or strongly agree).
The written feedback was effective in improving my leadership abilities.
The written feedback provided specific examples and areas for me to work on.
The oral feedback I received was effective in improving my leadership abilities.
The oral feedback provided specific examples and areas for me to work on.

Note: The study team assessed the reliability of scales using item responses from the second round of principal supervisor surveys, conducted in spring 2017. We measured reliability using Cronbach's alpha, a commonly accepted statistic for determining scale reliability. Coefficients above 0.90 are considered to have excellent internal consistency, those above 0.80 have good internal consistency, and those at 0.70 above are considered acceptable (DeVellis 2012).

Measures of the implementation of PSI components. We could only examine how measures of PSI implementation related to PSI effects when the implementation measures varied across supervisors within districts. In a correlational analysis using measures that only vary across districts, we would not be able to disentangle the effects of these measures from other district-level factors. Therefore, we developed measures of PSI implementation for the three PSI components that could vary across supervisors within districts: (1) reducing principal supervisors' span of control, (2) training supervisors and developing their capacity to support principals, and (3) strengthening central office structures to support and sustain changes in the principal supervisor's role. Two other components—revising the principal supervisors' job description to focus on instructional leadership and developing systems to identify and train new supervisors (succession planning)—did not vary within districts.

To construct scales, we conducted factor analyses using the survey data from spring 2017, midway through the initiative, when implementation was fully underway in all six districts.

Span of control. To measure the implementation of a reduced span of control, we used a single item from the principal supervisor survey: the supervisor's reported span of control.

Training. We used two principal supervisor-reported measures of the training provided as part of the PSI (Table A.3). To measure the reported quality of training provided, we constructed a scale using 19 items from the supervisor survey. The Cronbach's alpha (a measure of reliability ranging from 0 to 1) for the 19-item scale is high (0.98). Two main factors emerged from the exploratory factor analysis, but they were not conceptually distinct, so we combined these items into a single scale. To measure the extent to which the training provided to supervisors aligned with the PSI, we constructed a scale of 6 items from the supervisor survey that assessed the extent to which the training emphasized principal instructional leadership. These items aligned with one of the factors that emerged from the full set of 16 items related to the focus of supervisor training. The Cronbach's alpha is 0.87.

Central office. We constructed two measures of the quality of central office support (Table A.4). To measure the perceived quality of the central office in supporting supervisors' work with principals focused on teaching and learning, we constructed two measures—one based on supervisor reports and one based on principal reports. To measure quality of central office support as perceived by supervisors, we constructed a scale of 11 items from the supervisor survey (Cronbach's alpha = 0.90). To measure quality of central office support as perceived by principals, we constructed a scale of 11 items from the principal survey and then averaged the scale to the supervisor level (Cronbach's alpha = 0.90).

Supervisor time spent on instructional leadership (principal reported). The underlying theory of the PSI was that shifting supervisors' practices to focus more on instructional leadership would lead to improvements in principals' performance. To examine the extent to which supervisors prioritized instructional leadership in their time with principals, we constructed a measure of each supervisor's average time spent on instructional leadership activities with each principal. As our primary measure, we constructed a single measure of each supervisor's average estimated time spent on instructional leadership activities with each principal by multiplying three items from the principal survey—number of visits by supervisor at school, duration of visits by supervisor at school (in minutes), and percentage of the time spent on instructional leadership. We then averaged this measure this across the principals each supervisor oversees to construct a single measure for each supervisor.

To account for the possibility of measurement error in our measure of supervisor time spent on instructional leadership, given that principals' interpretations of the survey questions about length and frequency of principal meetings with supervisors may have varied, we also constructed an alternate measure of supervisor focus on instructional leadership. For this measure, we used a single item from the principal survey—the percentage of time principals reported a supervisor spending on instructional leadership.

Supervisor effectiveness (principal-reported). The PSI also sought to improve supervisors' effectiveness as instructional leaders in order to improve principals' performance. To measure supervisor effectiveness, we constructed three different measures, each based on principal reports averaged to the supervisor level (Table A.5). As a measure of principal-perceived overall effectiveness, we constructed a scale of 30 items about supervisor effectiveness from the principal survey (Cronbach's alpha = 0.98). We measured the extent to which supervisors implemented the specific practices that were the focus of the PSI by constructing a scale of 18 items from the principal survey about supervisors' implementation of these practices (Cronbach's alpha = 0.93). To measure the quality of supervisors' oral and written evaluation and feedback for principals, a major focus of the PSI, we constructed a scale of 4 items from the principal survey (Cronbach's alpha = 0.94).

The descriptive statistics for the measures of the implementation components, time spent on instructional leadership, and supervisor effectiveness are provided in Tables A.6 and A.7. Table A.6 provides the means and standard deviations of the measures, by year, and Table A.7 provides the correlations among the measures.

Table A.6. Means and standard deviations of the measures of PSI implementation, supervisors' practices, and supervisors' effectiveness

	2016		2017		2018		All years[a]	
Measure (scale unless otherwise noted)	*N*	M (SD)	*N*	M (SD)	*N*	M (SD)	*N*	M (SD)
Supervisor's span of control	517	15.6 (5.4)	506	13.4 (2.6)	544	13.5 (3.0)	563	14.2 (3.1)*
Supervisor training quality	211	3.9 (0.6)	416	3.7 (0.6)	444	3.7 (0.5)	519	3.7 (0.5)*
Alignment of supervisor training with PSI goals	211	2.1 (0.5)	416	2.2 (0.5)	444	2.1 (0.4)	519	2.1 (0.4)*
Quality of central office support according to supervisors	517	2.7 (0.8)	506	3.0 (0.7)	544	3.1 (0.7)	563	2.9 (0.6)*
Quality of central office support according to principals	528	3.0 (0.8)	503	3.2 (0.8)	554	3.3 (0.8)	564	3.2 (0.7)*
Supervisor's time spent on instructional leadership in past 3 months, as reported by principals (minutes)	513	208.9 (246.9)	479	213.1 (285.1)	537	211.3 (279.0)	563	205.4 (194.1)*
Percentage of supervisor's time spent on instructional leadership in past 3 months	531	61.7 (24.2)	510	54.7 (25.1)	559	54.0 (24.9)	564	56.3 (19.6)*
Principals' ratings of supervisor's effectiveness	529	3.9 (0.8)	507	3.9 (0.8)	558	4.1 (0.8)	564	4.0 (0.6)*
Supervisor's implementation of PSI's focal practices	531	3.4 (0.7)	510	3.5 (0.8)	561	3.6 (0.8)	564	3.5 (0.6)*
Principals' perceptions of evaluation feedback	487	4.0 (0.8)	476	4.1 (0.8)	526	4.2 (0.7)	562	4.1 (0.6)*

Source: Principal and supervisor survey data from 2016, 2017, and 2018.

[a] Measures are averaged across the three years in which principals and supervisors were surveyed. These values represent those used for the correlational analyses in the cumulative model.

* Significant difference across all years at the .05 level, two-tailed F-test.

N = number of schools; M = mean; SD = standard deviation.

Table A.7. Correlations among the measures of PSI implementation, supervisors' practices, and supervisors' effectiveness

	Correlation coefficients									
Measure	1	2	3	4	5	6	7	8	9	10
1. Span of control	1.00									
2. Training quality	0.00	1.00								
3. Training alignment	–0.17*	0.57*	1.00							
4. Central office support (S)	0.02	0.35*	0.13*	1.00						
5. Central office support (P)	–0.01	0.12*	0.06	0.32*	1.00					
6. Minutes spent on instructional leadership	–0.20*	–0.09*	–0.13*	0.03	0.14*	1.00				
7. Percentage of time spend on instructional leadership	0.05	0.06*	0.04	0.08*	0.17*	0.41*	1.00			
8. Supervisor's effectiveness	0.14*	0.10*	0.02	0.16*	0.28*	0.18*	0.29*	1.00		
9. PSI's focal practices for supervisors	0.02	0.06	–0.01	0.10*	0.25*	0.26*	0.32*	0.77*	1.00	
10. Evaluation feedback	0.09*	0.08*	0.06*	0.14*	0.23*	0.17*	0.25*	0.79*	0.65*	1.00

Source: Principal and supervisor survey data from 2016, 2017, and 2018.

Notes: The sample includes 1,071 schools.

Implementation measures are averaged across the three years in which principals and supervisors were surveyed: 2015–2016, 2016–2017, and 2017–2018.

* Significantly different from zero at the .05 level, two-tailed test.

IL = instructional learning; P = principal-reported measure; S = supervisor-reported measure.

6. Analysis of survey data

We included respondents in the sample for the survey data analysis if they met the following criteria:

1. Current principal supervisor or principal in a PSI district in that school year, according to district rosters
2. Employed in the role (principal or supervisor) at the time of survey administration
3. Supervisor or principal of a school that was part of the PSI

Some districts employed principals and principal supervisors who did not participate in the PSI, including supervisors and principals who worked in special or alternative schools that did not report to the same central office department as the one that housed most supervisors. We excluded these respondents from the survey analysis.

All analyses of the survey data are unweighted, giving equal weight to each respondent rather than weighting each district equally.

To examine the correlation between supervisors' span of control and their emphasis on key practices, we regressed span of control on the measures of emphasis on each of the four sets of key practices (data use, instruction, feedback, and classroom visits) for the two principals each supervisor reported on in the final round of the survey. We controlled for (1) whether the principal was experienced in a high-achieving school or inexperienced in a low-achieving school, (2) the supervisors' years of experience, and (3) district. We then used the coefficients from the regression analysis to form a fitted line representing the relationship between span of control and each set of key practices.

A representative estimation model for this analysis is:

$$(1) \qquad I_s = B_0 + B_1^{'} D_s + B_2^{'} X_g + d_i + u_g$$

where I_s represents the supervisors' emphasis on key practices for school *s*, D_s indicates whether school *s* is a high achieving school with an experience principal, X_g represents characteristics of supervisor *g* including their of experience as a supervisor and the number of schools within their span of control, d_i is a fixed-effect for each district participating in the PSI, and u_g is a random error term, clustered at the supervisor level (where supervisor assignments were identified based on the 2017–2018 school year). The estimated coefficient β_1 indicated the association of high- or low-identified schools with the supervisors' emphasis on key practices within that school.

B. Estimating the PSI's effects on teachers' perceptions of principals' performance

To estimate the PSI's effects on teachers' perceptions of principals' performance, we used a propensity score matching approach to compare teachers' perceptions of principals' performance in the PSI schools to those of principals in a similar set of non-PSI schools. We measured teachers' perceptions of principals' performance using teachers' reports of principals' performance from the VAL-ED. In this section we provide additional details to supplement the discussion of our approach in Chapter II.

1. Measuring teachers' perceptions of principals' performance

In this section we provide more details about the VAL-ED data used to measure teachers' perceptions of principals' performance.

The structure and content of the VAL-ED. Principal supervisors, principals, and teachers all take the same VAL-ED survey, which contains 72 items, each on a 5-point scale. Principal supervisors and principals take the full survey; teachers are randomly assigned to take one of two halves to reduce response time with little reduction in reliability (Porter et al. 2008). For each set of raters (principal supervisor, principal, and teachers), total scores for the principal are calculated as the mean response across the relevant survey items among that set of raters (Porter et al. 2010a). The VAL-ED items can also be grouped into subscales by organizing them in two different ways—either as six "core components" of 12 items each or six "key processes" of 12 items each. We considered estimating impacts on these subscales but determined that they were too highly correlated to measure distinct constructs, with correlations greater than 0.90 for all pairs of subscales.

Reliability and validity of the VAL-ED. Porter et al. (2010a) investigated the reliability and validity of the VAL-ED and found that the total score had an internal consistency reliability of 0.98 or above for supervisor, principal, and teacher reports. Covay Minor et al. (2014) examined the known group validity of the VAL-ED and found that VAL-ED scores were able to identify statistically significant differences between the top- and bottom-performing 20 percent of principals as identified by their superintendents. Porter et al. (2014) examined the correlation of principals' VAL-ED scores with their schools' estimated value added in promoting student achievement growth and found a significant correlation in some but not all model specifications.

Sufficient variation in VAL-ED scores. It is possible that teacher ratings might not exhibit sufficient variation if teachers tend to give their principals average scores (for example, because they have experienced few other principals to which they can compare their principal). Additionally, if the VAL-ED scores bunched at either the maximum or the minimum possible value, then it would suggest that the scores did not adequately differentiate between different types of principals due to a restriction on range. For this reason, we explored whether the VAL-ED scores exhibited sufficient variation to serve as a meaningful outcome. Figure A.1 shows that the VAL-ED scores followed a bell-shaped curve without any bunching at the tails, suggesting that they exhibited sufficient variation in our sample.

Figure A.1. Distribution of VAL-ED scores by year

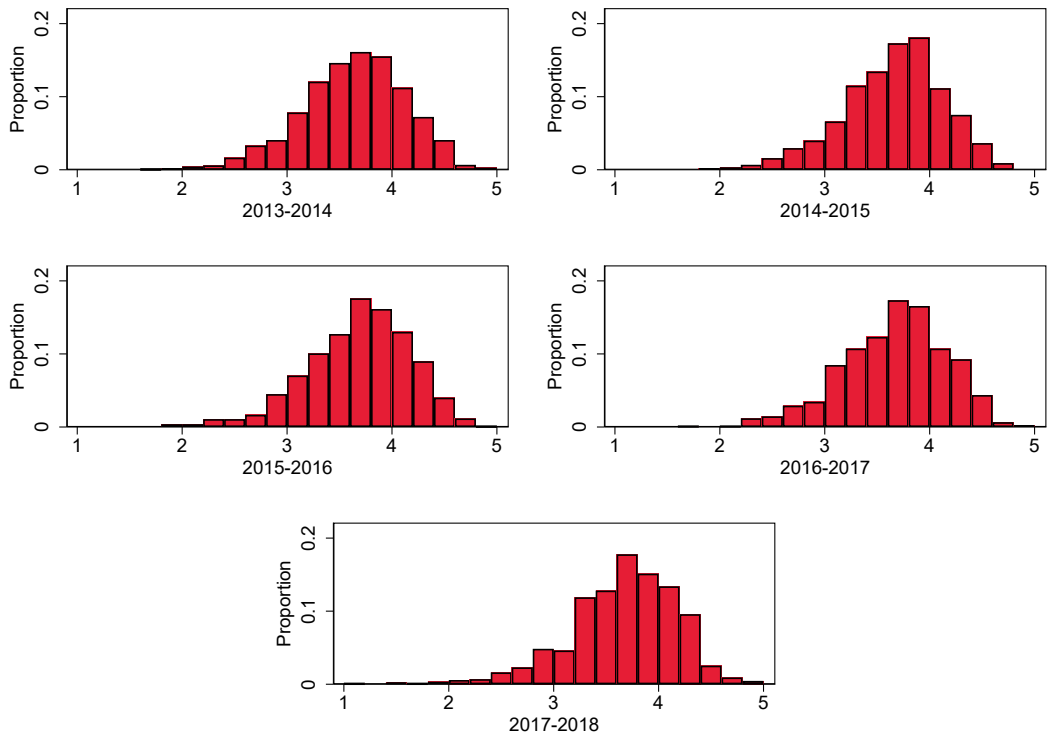


Figure reads: In each year, VAL-ED scores for schools in the study sample followed a bell-shaped curve without any bunching at the tails, suggesting that they exhibited sufficient variation.

Sources: VAL-ED survey data (2013–2014 to 2017–2018 school years).

Note: The figures show the distribution of the mean of teacher's reports on the VAL-ED score by year for the sample of all available schools.

Correlations between VAL-ED scores by respondent types. We calculated the correlation between the scores from each set of raters on the VAL-ED to examine whether teacher reports were related to principal and supervisor reports (Table A.8). The correlations ranged from 0.25 to 0.35, suggesting that the scores from the three raters are related but that the teacher reports do not fully proxy either the principal or supervisor reports.

Table A.8. Correlations among VAL-ED scores by respondent type

Respondent #1	Respondent #2	Correlation
Teacher	Principal	0.25
Teacher	Supervisor	0.35
Principal	Supervisor	0.32

Sources: VAL-ED survey data (2013–2014 school year).

Note: The table shows the correlation of VAL-ED scores by respondent type for the sample of all available schools at baseline, before the PSI began.

The possible consequences of measurement error in the VAL-ED scores. Measurement error arises in scales like the VAL-ED when the measure does not completely capture the underlying construct (in this case, teachers' perceptions of principals' performance). For example, measurement error might arise if some teachers respond randomly to some questions rather than carefully considering their responses. Measurement error in the VAL-ED score could potentially affect the estimates of the effect of the PSI in two ways, depending on whether the error arises in the follow-up or baseline scores:

1. **Measurement error in the follow-up VAL-ED scores could reduce statistical power.** Outcomes that suffer from measurement error introduce noise into estimation, thereby making it more challenging to attribute variation in the outcome to a particular variable, such as participation in the PSI. As a result, this noise tends to inflate estimates of the standard errors. Our power calculations accounted for this possibility because they were based on the estimated standard errors, which incorporated this type of noise. Based on these calculations, we determined that we had adequate power to detect meaningful effects.
2. **Measurement error in the baseline VAL-ED score could potentially bias estimated effects.** In propensity score matching, measurement error in the matching variables can lead to a bias in the estimated effect because the matching variables may no longer adequately account for selection into the PSI districts (Steiner et al. 2011). Given that the baseline VAL-ED score is a key matching variable, measurement error in the VAL-ED score could potentially bias the effect estimates. However, any such bias is likely to be minimal because the VAL-ED has an internal consistency reliability above 0.98 (Porter et al. 2010a), suggesting it has low levels of measurement error.

Response rates. Nonresponse on the VAL-ED survey could lead to a bias in the estimated effect of the PSI. If teachers in the PSI districts were more (or less) likely to respond to the VAL-ED survey than teachers in non-PSI districts, changes in the VAL-ED score might not reflect true changes in teachers' perceptions of principals' performance. For example, such a bias could arise if PSI districts encouraged teachers who have a more positive view on their performance to respond to the survey. Then, relative to comparison schools, the VAL-ED score in PSI schools might appear higher even if the principal's performance did not improve. To shed light on this possibility, we considered nonresponse rates in PSI and comparison schools over time (Table A.9). Relative to the comparison districts, the PSI districts had somewhat higher response rates in each year. Nevertheless, for all but the last year, the response rates were similar between the PSI and comparison districts, suggesting little room for bias. In addition, all our analyses control for baseline response rates, further mitigating any potential biases.

Table A.9. Average teacher response rates on VAL-ED by year

Sample	Average teacher response rate on VAL-ED				
	2013–2014	2014–2015	2015–2016	2016–2017	2017–2018
All schools	**76**	**84**	**80**	**84**	**79**
All schools in PSI districts in the analysis sample	80	87	81	85	87
All potential comparison schools	58	81	77	73	56
Matched comparison schools	81	82	76	76	58
Difference between PSI schools in the analysis sample and matched comparison schools	–1	5	5	9	29

Sources: VAL-ED survey data (2013–2014 to 2017–2018 school years).

Note: The baseline sample (2013–2014) includes schools from the 2017–2018 analysis sample. The samples for the remaining years are restricted to those cases that have a baseline VAL-ED score, as in our main analyses. The sample excludes districts that are in The Wallace Foundation's Principal Pipeline Initiative and two exemplar districts that received smaller PSI grants but are not part of the evaluation. Because all schools in PSI districts were classified as either urban or suburban, we also restrict the comparison schools to those classified as urban or suburban. To calculate the overall averages, we first calculate the average response rate for each school. We then calculate the average of the school-level response rates across schools, potentially giving different weights to different schools. The sample of all schools weights each schools equally and includes all PSI and comparison school that enter the analysis. The PSI and matched comparison schools samples are weighted as they are in the effects analyses. The baseline sample (2013–2014) includes all schools that enter the analysis in the 2017–2018 year and uses the corresponding weights from that analysis year.

2. Constructing a matched comparison group for the analysis of the PSI's effects on teachers' perceptions of principals' performance

In this section, we provide more details about the features of our propensity score matching approach for construction of a matched comparison group.

Approach to estimating propensity scores. A first step to implementing the matching procedure is to estimate for the propensity scores for both the PSI and non-PSI schools—that is, the probability that each school participated in the PSI as a function of their baseline characteristics. The propensity score summarizes the extent to which two schools are similar before the start of the PSI. For the PSI schools and potential comparison schools (the universe of schools that implemented the VAL-ED in the 2013–2014 school year and subsequent years of the study), we estimated a logit model of the likelihood of a school being in a PSI district as a function of relevant school-level factors, measured during or before the baseline (2013–2014) school year (see Table A.10 for a list of these variables). We then used the results of this model to predict a propensity score reflecting the probability that each school was in a PSI district.

Matching approach. To form a matched comparison group, we used a propensity score matching procedure called kernel matching. In this procedure, each observation in the treatment group is matched to a weighted average of similar observations in the comparison group, thereby differing from methods that match each treatment observation to a single comparison observation and methods that rely on linear regressions to adjust for differences. By using more information,

kernel weighting approaches reduce asymptotic mean squared error compared with methods that use a single matched observation (Smith and Todd 2005).

Weighting of comparison schools. Kernel matching uses the estimated propensity scores to measure the degree of similarity between PSI schools and potential comparison schools. In particular, it identified comparison schools as matches to a specific PSI school if their propensity scores were within a maximum distance (bandwidth) from the PSI school's propensity score. Among those matched comparison schools, those with a propensity score that was closer to the PSI school's propensity score received a larger weight. The kernel is the function that determines the weight. We used weights based on the Epanechnikov kernel, a common approach that maximizes efficiency when estimating distributions. We also tested the sensitivity of the estimates to the bi-weight kernel, another commonly used approach (Appendix Table B.1).

Bandwidth selection. We selected the bandwidth for the kernel using an approach known as least squares cross-validation (Hall et al. 2004; Frölich 2004). In this approach, data from comparison schools was used to select the bandwidth that minimized the sum of the squared differences between the true outcome for each school and the expected outcome for the school given the school's propensity score. Each school's expected outcome was estimated using data from all schools in the comparison group aside from that school. We assessed the sensitivity to the bandwidth selection by using one-third and three times the cross-validation bandwidth (Huber et al. 2013), as well as Silverman's rule of thumb (Silverman 1986), which suggests picking the optimal bandwidth assuming that the data follow a known distribution with a variance equal to the sample variance.

Ensuring common support. One key assumption of all matching estimators is that the treatment and comparison groups have similar propensity scores—a condition referred to as common support. A common support helps ensure that all schools can be matched to other schools with similar propensity scores. To ensure common support, we trimmed the data by dropping observations for which there is not a comparable match. In particular, for our main estimates, we removed any observation in the treatment group that had a propensity score that was greater than the maximum in the comparison group.

Balancing tests. As described in the text, we conducted balancing tests to assess the difference in baseline characteristics between the PSI schools and matched comparison schools. We assess the quality of the matches by examining whether the PSI schools and matched comparison schools are well balanced in that they have similar characteristics at baseline, including those used to construct the propensity score. Following the recommendations of Rubin (2001), as summarized by Stuart (2010), we examined two sets of statistics to assess balance: (1) the standardized difference in mean values of the baseline variables and propensity scores between treatment and comparison groups calculated using the kernel weights and (2) the ratio of the variances of the propensity scores in the treatment and comparison groups. We considered the groups well balanced if the absolute value of the standardized difference in means is less than 0.25 (that is, the means differ by less than one-quarter of a standard deviation) and the variance ratio of the propensity scores is between 0.5 and 2.0 (the larger value of the variance no more than twice the smaller value). When our initial model was not well balanced, we adjusted the propensity score model (for instance, adding higher-order terms of particular variables) to improve balance. Table A.10 presents these results. As discussed in the text, all but two of the

effect size estimates are less than the threshold for large differences. The two differences that do not meet this criteria are not statistically significant. Additionally, the ratio of the variance of propensity scores falls within the acceptable range of all years.

Table A.10. Balancing tests on baseline covariates and propensity scores for analytic samples

	Effect size differences between PSI and comparison matches for analytic samples (unless otherwise indicated)			
Variable	2014–2015	2015–2016	2016–2017	2017–2018
School characteristics				
School size, level, and locale				
Total enrollment	0.14	–0.11	–0.12	0.07
Number of full-time equivalent teachers	0.08	–0.18	–0.21	0.06
Secondary school	–0.07	–0.02	–0.25	–0.39
Urban locale	–0.17	0.17	0.03	0.01
School demographic composition				
Percentage black, non-Hispanic	–0.05	0.23	0.10	–0.22
Percentage white, non-Hispanic	0.06	–0.09	0.10	0.14
Percentage Hispanic	–0.05	–0.16	–0.16	0.12
Percentage eligible for free or reduced-price lunch	0.07	0.14	0.03	–0.15
Student achievement				
School math and reading proficiency rates				
Math, baseline year	0.05	–0.12	–0.08	0.11
Math, one year prior	0.03	–0.11	0.02	0.18
Math, two years prior	0.10	–0.01	0.07	0.19
Reading, baseline year	0.02	–0.11	0.00	0.07
Reading, one year prior	–0.11	–0.16	0.01	0.00
Reading, two years prior	–0.01	–0.06	0.03	0.09
VAL-ED scores				
VAL-ED scores	–0.01	–0.03	–0.20	0.09
Percentage of teachers responding to the VAL-ED survey	0.10	–0.16	–0.26	–0.04
Propensity scores				
Effect size difference	0.00	0.00	0.00	0.00
Ratio of variances of propensity scores	1.00	1.00	1.02	1.00

Sources: VAL-ED survey data (2013–2014 school year), Common Core of Data (2013–2014 school year), and EdFacts data (2011–2012 to 2013–2014 school years).

Note: All variables used for matching are from the baseline (2013–2014) school year unless otherwise indicated. The mean of the comparison schools was calculated using the weights derived from the propensity score matching procedure. The sample excludes districts that are in The Wallace Foundation's Principal Pipeline Initiative and two exemplar districts that received smaller PSI grants but are not part of the evaluation. Because all schools in PSI districts were classified as either urban or suburban, we also restrict the comparison schools to those classified as urban or suburban.

3. Estimating of the PSI's effects on teachers' perceptions of principals' performance

In this section, we provide more details about how we estimated the effects of the PSI.

Approach to calculating standard errors. For all models, we estimated standard errors using bootstrap methods that account for both the estimation error in the propensity scores and the correlation of outcomes within districts. Propensity score matching involves two-stages: (1) estimating propensity scores, and (2) forming weighted differences (potentially adjusted for covariates) between the treatment and comparison groups. In general, the estimated standard

errors from the second stage will not reflect the fact that the propensity scores are estimates that add noise to the estimated effect. Bootstrap methods account for this estimation error. Another issue is that outcomes might be correlated within districts (the level of assignment to PSI), which could lead to biased standard errors. We accounted for clustering in our bootstrap approach by resampling (with replacement) data at the district-level, rather than the school-level (Cameron and Miller 2015). To account for estimation error in the propensity scores, we recalculated both the propensity scores and the treatment effects in each bootstrap replication. We also trimmed the data in each bootstrap replication, thereby ensuring that each sample satisfies the common support condition.

Approach to calculating *p*-values. Additionally, standard tests of statistical significance can over-reject the null hypothesis when there are relatively few clusters, as is the case for the 2017–2017 and 2017–2018 analysis samples. To account for this issue, we calculated *p*-values using the pairs cluster bootstrap-t procedure, which has been shown to perform relatively well with few clusters (Cameron et al. 2008).

C. Examining how teachers' perceptions of principals' performance and the PSI's effects vary across districts and supervisors

As described in Chapter II, we estimated a measure known as the intracluster correlation to measure the percentage of the overall variation in the outcome (teachers' perceptions of principals' performance or school-level PSI effects) that is accounted for by differences between supervisors and between districts. To estimate the percentage of variance attributed to supervisors and districts, we first calculated the ICC for schools:

$$ICC_{\text{school}} = \frac{\sigma^2_{\text{school}}}{\sigma^2_{\text{supervisor}} + \sigma^2_{\text{district}} + \sigma^2_{\text{school}} + \sigma^2_{\text{year}}} \quad (2)$$

where $\sigma^2_{\text{supervisor}}$ is the variance between supervisors, $\sigma^2_{\text{district}}$ is the variance between districts, σ^2_{school} is the variance between individual schools, and σ^2_{year} is the variance between years. The school ICC expresses the similarity in performance or effects across years within the same school.

For this analysis, we assumed that variation across years within the same school is due to measurement error, and we adjusted our estimate of the overall variation to remove this variation. We consider the variation across years within the same school to be an upper bound of possible measurement error, since this variation could in part reflect actual changes in performance (or treatment effects). Using our adjusted estimate of the overall variation in the outcome, we then calculated the supervisor and district ICCs as follows:

$$ICC_{\text{supervisor}} = \frac{\sigma^2_{\text{supervisor}}}{\sigma^2_{\text{supervisor}} + \sigma^2_{\text{district}} + \sigma^2_{\text{school}}} \quad (3)$$

$$(4) \qquad ICC_{\text{district}} = \frac{\sigma^2_{\text{district}}}{\sigma^2_{\text{supervisor}} + \sigma^2_{\text{district}} + \sigma^2_{\text{school}}}$$

The supervisor ICC provides an upper bound on the percentage of variation attributed to principal supervisors, and the district ICC provides an upper bound on the percentage of variation attributed to districts for the error-adjusted measures.

D. Examining factors associated with the PSI's effects

Although the PSI had no effects on teachers' perceptions of principals' performance on average across all the schools in the initiative, effects varied considerably across individual schools (Figure A.2). Along with the variation in PSI implementation, this variation in effects across schools suggests that differences in how individual districts and supervisors implemented the PSI could relate to differences in the PSI's effects across schools.

Figure A.2. Distribution of estimated PSI effects (in VAL-ED units) across all schools in PSI districts in 2018

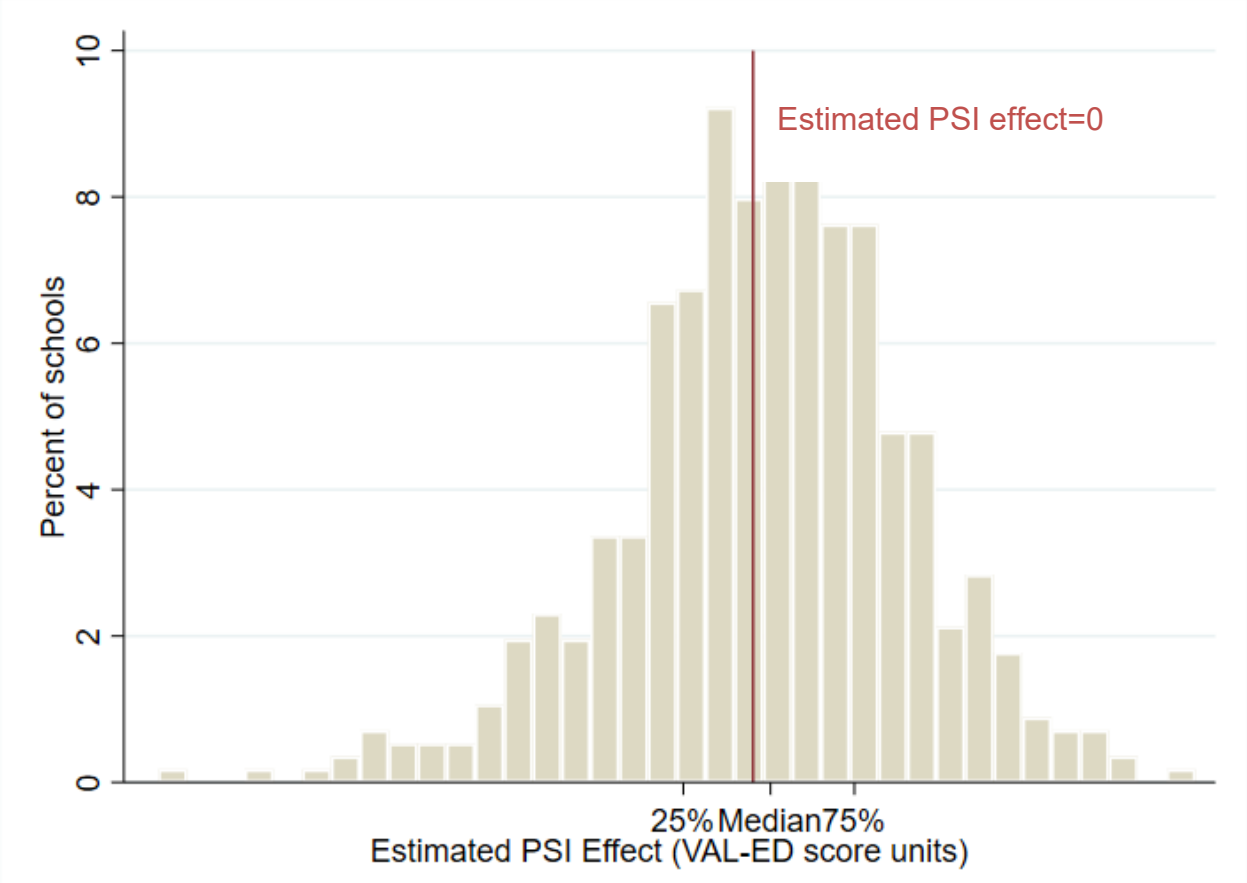


Figure reads: The estimated effects of the PSI on teachers' perceptions of principals' performance varied substantially across schools in PSI districts in 2018.

Source: VAL-ED survey data from 2018.

To examine the correlation between the PSI's effects on teachers' perceptions of principals' performance and the implementation measures described previously, we regressed estimated effects of the PSI for each school on each of the measures of PSI implementation, supervisors' practices, and supervisors' effectiveness shown in Table A.6. For measures of PSI implementation of individual components (span of control, training, and central office support), we first estimated bivariate regressions for each individual component measure. We had then planned to estimate a multivariate regression that included any of the individual component measures with a positive, statistically significant regression coefficient in the bivariate analysis. In practice, only one component measure (supervisors' span of control) was significantly associated with estimated effects of the PSI on teachers' perceptions of principals' performance.

A representative estimation model for this analysis is:

$$(5) \qquad I_s = \beta_0 + \beta_1' P_s + u_s$$

where I_s represents the effect of the PSI on teachers' perceptions of principals' performance in school s, P_s represents the values of a given set of implementation, practice, or effectiveness measures for the supervisor(s) overseeing school s, and u_s is a random error term, clustered at the supervisor level (where supervisor assignments were identified based on the 2017–2018 school year). We weighted each observation by the inverse of the variance of the effect estimate for that school to give more weight to more precisely estimated effects. The estimated coefficient β_1 indicated the association of each implementation measure with the PSI's effects on teachers' perceptions of principals' performance.

For our main model, I_s is estimated in the 2017–2018 school year and represents the cumulative effects of the PSI on teachers' perceptions of principals' performance over four years. Because the characteristics of implementation throughout the grant period may have contributed to these cumulative effects, we averaged the implementation measures across the 2015–2016, 2016–2017, and 2017–2018 school years (the three years of the PSI during which data were collected).

APPENDIX B.

SUPPLEMENTAL ANALYSES AND SENSITIVITY TESTS

This appendix provides results from supplemental analyses and sensitivity tests related to the PSI's effects on teachers' perceptions of principals' performance, the variation in teachers' perceptions of principals' performance and on PSI's effects, and the correlations between components of the PSI and PSI's effects on teachers' perceptions of principals' performance. These results provide additional documentation to support the findings presented in the main body of the report.

A. Supplemental analyses for estimating the PSI's effects on principals' performance

In this section, we present the results of sensitivity analyses that we conducted for our analysis of the PSI's effects on teachers' perceptions of principals' performance. We also present results from exploratory analyses that examine the PSI's effects on alternate measures of principals' performance: principal- and supervisor-reported VAL-ED scores as well as aggregate VAL-ED scores that combine reports from teachers, principals, and supervisors.

Across all nine sensitivity analyses and all years, point estimates were generally similar, and none of the estimated effects on teachers' perceptions of principals' performance were statistically significant (Table B.1).

- Sensitivity tests 1–3 show results from models that differ from the main model in terms of the bandwidth. We used one-third and three times the cross-validation bandwidth used for the main analysis (Huber et al. 2013), as well as Silverman's rule of thumb (Silverman 1986), which suggests picking the optimal bandwidth assuming that the data follow a known distribution with a variance equal to the sample variance.
- Sensitivity test 4 shows results from a model that uses the bi-weight kernel, a commonly used alternative to the Epanechnikov kernel.
- Sensitivity test 5 shows estimates of the average treatment effect, rather than the treatment-on-the treated effect. Our main analysis estimated the effect of PSI on the schools that participated in the program. The literature sometimes refers to this estimate as the effect of the treatment on the treated. However, we also estimated the effect of the PSI for the full sample of schools, including the schools that are not in PSI districts. This estimate is often referred to as the average treatment effect, and it represents the effect of PSI if all schools in the sample (both in PSI and comparison districts) had been in the initiative.[14] When estimating this parameter, we adjusted our trimming rules. In particular, we also removed any observation in the comparison group that had a propensity score that was lower than the lowest score in the treatment group.
- Sensitivity test 6 shows estimates from an approach that is equivalent to the main model but does not additionally control for covariates after estimating the kernel weights.

[14] The average treatment effect could differ from the effect of the treatment on the treated if the PSI has a different effect on different types of schools. For example, it is possible that the effect of the PSI would have been greater for schools outside of the six PSI districts because those schools might have been less likely to provide similar supports in the absence of the PSI. In that case, the average treatment effect would exceed the effect of the treatment on the treated.

- Sensitivity test 7 shows estimates using the main model but redefining the samples so that they are relative to the start of implementation, rather than the school year. In particular, this shift involves using the subsequent year of data for Baltimore, because Baltimore implemented the program a year after all other districts. For example, the "first year" sample includes 2015–2016 data for Baltimore and 2014–2015 data for all other samples.
- Sensitivity test 8 shows estimates using inverse probability weighting, rather than kernel matching. In this method, the treatment-on-the-treated parameter is the mean difference between the treatment group and the comparison group, where the schools in the comparison group are reweighted so that their characteristics match the distribution of the characteristics of schools in the treatment group. Inverse probability weighting has also been shown to perform well (Busso et al. 2014).
- Sensitivity test 9 shows estimates from an ordinary least squares model, in which the dependent variable is the VAL-ED score and the independent variables include a dummy variable for PSI participation and the set of covariates used in the kernel matching procedure. This specification represents a parametric alternative for accounting for baseline differences and offers more statistical power.

Table B.1. Estimated effects of PSI on teacher-reported VAL-ED scores by school year (sensitivity analyses)

	Mean VAL-ED score (teacher report)[a]					Number of schools		Number of districts	
Year	PSI	Comparison	Effect	SE	*p*-value	PSI	Comparison	PSI	Comparison
	Main model (kernel matching, Epanechnikov kernel, cross-validation bandwidth, treatment on the treated effect, additional controls for covariates)								
2014–2015	–0.01	–0.04	0.01	0.09	0.505	490	327	5	91
2015–2016	0.03	0.09	–0.05	0.06	0.575	628	347	6	69
2016–2017	0.05	0.00	–0.01	0.10	0.763	585	92	6	22
2017–2018	0.08	0.08	0.03	0.10	0.302	612	208	6	20
	Sensitivity test 1 (main model but with 1/3 times the cross-validation bandwidth)								
2014–2015	–0.01	–0.07	0.02	0.09	0.430	462	260	5	85
2015–2016	0.03	0.07	–0.04	0.06	0.547	550	290	6	59
2016–2017	0.03	–0.02	0.00	0.11	0.775	553	92	6	20
2017–2018	0.08	0.05	0.05	0.11	0.272	547	159	6	18
	Sensitivity test 2 (main model but with 3 times the cross-validation bandwidth)								
2014–2015	–0.01	–0.04	0.01	0.09	0.541	490	380	5	95
2015–2016	0.03	0.06	–0.04	0.06	0.519	628	394	6	80
2016–2017	0.05	–0.02	0.02	0.09	0.706	593	105	6	24
2017–2018	0.08	0.08	0.02	0.10	0.392	612	252	6	21
	Sensitivity test 3 (main model but with Silverman's rule of thumb bandwidth)								
2014–2015	–0.01	–0.01	–0.01	0.08	0.614	490	380	5	95
2015–2016	0.03	0.05	–0.03	0.06	0.489	628	394	6	80
2016–2017	0.05	–0.03	0.03	0.10	0.630	593	112	6	25
2017–2018	0.08	0.08	0.03	0.09	0.337	612	252	6	21
	Sensitivity test 4 (main model but with the bi-weight kernel)								
2014–2015	–0.01	–0.04	0.01	0.09	0.491	490	327	5	91
2015–2016	0.03	0.09	–0.05	0.06	0.573	628	354	6	71
2016–2017	0.05	0.00	–0.01	0.10	0.769	585	93	6	22
2017–2018	0.08	0.08	0.04	0.10	0.274	612	203	6	19
	Sensitivity test 5 (main model but with the average treatment effect)								
2014–2015	–0.05	–0.02	–0.02	0.06	0.632	490	327	5	91
2015–2016	0.01	0.07	–0.05	0.05	0.628	628	333	6	66
2016–2017	0.04	–0.01	0.02	0.09	0.714	582	99	6	22
2017–2018	0.08	0.06	0.03	0.13	0.330	612	204	6	20
	Sensitivity test 6 (main model but without additional controls for covariates)								
2014–2015	–0.01	–0.04	0.03	0.11	0.505	490	327	5	91
2015–2016	0.03	0.09	–0.06	0.08	0.765	628	347	6	69
2016–2017	0.05	0.00	0.04	0.14	0.716	585	92	6	22
2017–2018	0.08	0.08	0.00	0.16	0.472	612	208	6	20

	Mean VAL-ED score (teacher report)[a]					Number of schools		Number of districts	
Year	PSI	Comparison	Effect	SE	*p*-value	PSI	Comparison	PSI	Comparison
Sensitivity test 7 (main model but grouping samples by years since implementation, rather than school year)									
First year	0.01	–0.02	0.01	0.07	0.429	629	283	6	88
Second year	0.07	0.09	–0.02	0.06	0.505	625	362	6	74
Third year	0.05	0.00	–0.01	0.10	0.757	585	92	6	22
Fourth year	0.04	0.03	0.01	0.12	0.355	477	239	5	21
Sensitivity test 8 (inverse probability weighting, treatment on the treated effect)									
2014–2015	–0.01	–0.14	0.14	0.22	0.352	484	380	5	95
2015–2016	0.03	–0.04	0.07	0.20	0.409	628	394	6	80
2016–2017	0.05	–0.07	0.13	0.25	0.481	488	137	6	27
2017–2018	0.08	0.00	0.09	0.31	0.322	612	252	6	21
Sensitivity test 9 (ordinary least squares)									
2014–2015	–0.02	0.01	–0.03	0.04	0.484	498	380	5	95
2015–2016	–0.01	0.03	–0.05	0.04	0.496	628	394	6	80
2016–2017	0.01	0.00	0.01	0.06	0.562	618	137	6	27
2017–2018	–0.01	0.00	–0.01	0.07	0.424	614	252	6	21

Sources: VAL-ED survey data (2013–2014 to 2017–2018 school years), Common Core of Data (2013–2014 school year), and EdFacts data (2011–2012 to 2013–2014 school years).

Note: The sample excludes districts that are in The Wallace Foundation's Principal Pipeline Initiative and two exemplar districts that received smaller PSI grants but are not part of the evaluation. Because all schools in PSI districts were classified as either urban or suburban, we also restrict the comparison schools to those classified as urban or suburban.

[a] For the kernel-matching specifications that control for covariates, the control mean is the weighted average among comparison schools in the estimation sample, and the treatment mean is the weighted control mean plus the estimated effect, where the weights are based on the estimation routine. For the kernel-matching specification that does not control for covariates and the inverse probability weighting, the means reflect the weighted average of outcomes in the PSI and comparison schools, where the weights are based on the estimation routine. For the ordinary least squares specification, the control mean is the average among comparison schools in the estimation sample and the treatment mean is the control mean plus the estimated effect. All standard errors are calculated using a bootstrap procedure where districts are resampled with replacement. None of the effects are statistically significant at the 5 percent level.

We also conducted exploratory analyses to examine how estimates of the PSI's effects might have differed if we measured principals' performance using principal- and supervisor-reported VAL-ED scores or aggregate VAL-ED scores that combine reports from teachers, principals, and supervisors. In these exploratory analyses, we followed the same approach as we did in the main analyses of teacher-reported VAL-ED scores, except that we used these alternative VAL-ED scores as the outcome and included baseline measures of the corresponding scores in the propensity score model instead of teacher-reported scores. As discussed in Chapter II, the main analyses focused on teacher-reported VAL-ED scores due to concerns that measures based on principals' and supervisors' ratings could be biased. Such bias might arise because the PSI might have affected not only the true performance of principals, but also *how* supervisors rated principals and how principals rated themselves. These exploratory analyses show that the PSI did not affect principals' performance as measured by principal- and supervisor-reported VAL-ED scores or aggregate scores (Table B.2). Even though these findings are consistent with our main results, we suggest interpreting them with caution due to the potential biases in principals' and supervisors' reports.

Table B.2. Estimated effects of PSI on aggregate and teacher-, principal-, and supervisor-reported VAL-ED scores by school year

	Mean VAL-ED score[a]				
Year	PSI	Comparison	Effect	Standard error	*p*-value
	Teacher-reported VAL-ED scores (main analysis)				
2014–2015	-0.01	-0.04	0.01	0.09	0.505
2015–2016	0.03	0.09	-0.05	0.06	0.575
2016–2017	0.05	0.00	-0.01	0.10	0.763
2017–2018	0.08	0.08	0.03	0.10	0.302
	Principal-reported VAL-ED scores				
2014–2015	0.07	0.14	-0.13	0.10	0.640
2015–2016	0.08	0.22	-0.10	0.10	0.612
2016–2017	0.14	0.05	0.03	0.13	0.579
2017–2018	0.14	0.16	-0.02	0.20	0.508
	Supervisor-reported VAL-ED scores				
2014–2015	0.06	0.10	-0.14	0.17	0.431
2015–2016	0.12	0.07	-0.22	0.15	0.316
2016–2017	-0.03	-0.25	-0.04	0.15	0.403
2017–2018	0.14	-0.01	-0.01	0.19	0.301
	Aggregate VAL-ED scores based on teacher, principal, and supervisor reports				
2014–2015	0.05	0.06	-0.07	0.08	0.530
2015–2016	0.08	0.10	-0.08	0.08	0.477
2016–2017	0.10	0.03	0.03	0.09	0.582
2017–2018	0.14	0.11	0.01	0.12	0.229

Sources: VAL-ED survey data (2013–2014 to 2017–2018 school years), Common Core of Data (2013–2014 school year), and EdFacts data (2011–2012 to 2013–2014 school years).

Note: The sample excludes districts that are in The Wallace Foundation's Principal Pipeline Initiative and two exemplar districts that received smaller PSI grants but are not part of the evaluation. Because all schools in PSI districts were classified as either urban or suburban, we also restrict the comparison schools to those classified as urban or suburban.

[a] For the aggregate and principal- and supervisor-reported VAL-ED scores, we use the same estimation approach as in the main analysis. All standard errors are calculated using a bootstrap procedure where districts are resampled with replacement. None of the effects are statistically significant at the 5 percent level.

B. Supplemental analyses for examining how teachers' perceptions of principals' performance and the PSI's effects vary across districts and supervisors

In Chapter VI, our analysis of variation in teachers' perceptions of principals' performance and the PSI's effects focused just on the PSI districts because we only had information on principals' supervisors and PSI effects in those districts. However, we also investigated whether the percentage of variation in teachers' perceptions of principals' performance accounted for by districts was similar outside of the PSI districts, in the comparison schools (Table B.3). The percentage of variation in teachers' perceptions of principals' performance accounted for by districts was similar between PSI and comparison schools.

Table B.3. Proportion of variation in teachers' perceptions of principals' performance accounted for by districts in PSI schools and comparison schools (intracluster correlation coefficients)

Group	Number of schools	Number of districts	Proportion of variation accounted for by districts
PSI schools	637	6	0.10
Comparison schools	690	124	0.17

Sources: VAL-ED survey data from 2015, 2016, 2017, and 2018.

Note: Comparison schools are those matched to treatment schools each year for estimating treatment effects. The intracluster correlation coefficient is the proportion of total variation in an outcome that is accounted for by differences between clusters (in this case, between districts). The ICCs are adjusted for measurement error by subtracting out the proportion of variation accounted for by differences between years within the same school.

C. Supplemental analyses for examining factors associated with the PSI's effects

To assess whether findings from our analysis of factors associated with the PSI's effects were robust to the choice of estimation model and sample, we conducted two sensitivity analyses.

Our main model related the implementation, instructional leadership time, and effectiveness measures as averaged across the 2015–2016, 2016–2017, and 2017–2018 school years to the cumulative effects of the PSI in the final year of the initiative (Table B.4). To account for the possibility that implementation factors in each year influenced the PSI's effects in that year, we also estimated a concurrent model that pooled data from all three years in which we measured implementation and related the implementation measures in each year to the PSI's cumulative effects as of that year. Findings from the concurrent model were similar to the findings from our main model (Table B.5).

Table B.4. Main analysis: relationship between implementation measures and PSI effects (bivariate regressions)

Implementation measure (scale unless otherwise indicated)	Sample size	Relationship between implementation measure and PSI effects (regression coefficient)	Standard error	*p*-value
Supervisor's span of control (number of principals overseen)	563	0.02*	0.01	0.01
Supervisor training quality	519	-0.04	0.05	0.49
Alignment of supervisor training with PSI goals	519	-0.03	0.06	0.63
Quality of central office support according to supervisors	563	-0.01	0.04	0.77
Quality of central office support according to principals	564	0.03	0.03	0.30
Supervisor's time spent on instructional leadership in past 3 months, as reported by principals (minutes)	563	<0.01	<0.01	0.30
Percentage of supervisor's time spent on instructional leadership in past 3 months	564	<0.01*	<0.01	0.04
Principals' ratings of supervisor's effectiveness	564	0.11*	0.04	0.01
Supervisor's implementation of PSI's focal practices	564	0.11*	0.04	0.00
Principals' perceptions of evaluation feedback	562	0.09*	0.04	0.02

Sources: Principal and supervisor survey data from 2016, 2017, and 2018 and VAL-ED survey data from 2018.

Note: Implementation measures are averaged across the three years in which principals and supervisors were surveyed: 2015–2016, 2016–2017, and 2017–2018.

For all regression models, the dependent variable is the PSI effects on principals' VAL-ED scores in 2018.

*Significantly different from zero at the .05 level, two-tailed test.

Table B.5. Sensitivity analysis: relationship between implementation measures and PSI effects in each year of PSI implementation (bivariate regressions from the concurrent model)

Implementation measure (scale unless otherwise indicated)	Sample size	Relationship between implementation measure and PSI effects (regression coefficient)	Standard error	*p*-value
Supervisor's span of control (number of principals overseen)	624	0.01	0.01	0.19
Supervisor training quality	557	-0.08*	0.03	0.00
Alignment of supervisor training with PSI goals	557	-0.11*	0.04	0.01
Quality of central office support according to supervisors	624	<0.01	0.02	0.64
Quality of central office support according to principals	621	0.04	0.02	0.08
Supervisor's time spent on instructional leadership in past 3 months, as reported by principals (minutes)	621	<0.01	<0.01	0.61
Percentage of supervisor's time spent on instructional leadership in past 3 months	624	<0.01	<0.01	0.08
Principals' ratings of supervisor's effectiveness	623	0.05 *	0.03	0.04
Supervisor's implementation of PSI's focal practices	624	0.05	0.03	0.06
Principals' perceptions of evaluation feedback	616	0.06*	0.02	0.01

Sources: Principal and supervisor survey data and VAL-ED survey data from 2016, 2017, and 2018.

Notes: Implementation measures from each year are regressed on the PSI effect estimate from the same year.
For all regression models, the dependent variable is the PSI effects on principals' VAL-ED score.
N = number of schools included in the regression analysis. SE = standard error. β = regression coefficient.

*Significantly different from zero at the .05 level, two-tailed test.

The sample for our main analyses included schools with implementation data from any of the three years during which the supervisor and principal surveys were administered. As a sensitivity test, we restricted our sample only to schools for which we collected implementation data in each of the three years of data collection. Findings were similar to those from our main model, in which between 14 and 30 percent of schools in the sample were missing implementation data in one or two of the three years, for most implementation measures (Table B.6).

Table B.6. Sensitivity analysis: Relationship between implementation measures and PSI effects for subsample of schools that had data in all three years (bivariate regressions)

Implementation measure (scale unless otherwise indicated)	Sample size	Relationship between implementation measure and PSI effects (regression coefficient)	Standard error	*p*-value
Supervisor's span of control (number of principals overseen)	458	0.01*	0.01	0.04
Supervisor training quality	170	0.06	0.08	0.50
Alignment of supervisor training with PSI goals	170	0.08	0.09	0.38
Quality of central office support according to supervisors	458	0.01	0.05	0.84
Quality of central office support according to principals	474	0.02	0.03	0.53
Supervisor's time spent on instructional leadership in past 3 months, as reported by principals (minutes)	427	<0.01	<0.01	0.34
Percentage of supervisor's time spent on instructional leadership in past 3 months	483	<0.01	<0.01	0.06
Principals' ratings of supervisor's effectiveness	479	0.11*	0.04	0.01
Supervisor's implementation of PSI's focal practices	484	0.11*	0.04	0.01
Principals' perceptions of evaluation feedback	391	0.02	0.05	0.71

Sources: Principal and supervisor survey data from 2016, 2017, and 2018 and VAL-ED survey data from 2018.

Notes: Implementation measures are averaged across the three years in which principals and supervisors were surveyed: 2015–2016, 2016–2017, and 2017–2018.

For all regression models, the dependent variable is the PSI effects on principals' VAL-ED scores in 2018.

*Significantly different from zero at the .05 level, two-tailed test.

Commissioned by: