REPORT

PART B

Impact Evaluation of Support for Principals: OMB Data Collection Package

June X, 2015

`

CONTENTS

TABLES

PART B. SUPPORTING STATEMENT FOR PAPERWORK
REDUCTION ACT SUBMISSION

This Office of Management and Budget (OMB) package requests clearance for data collection activities to support the Impact Evaluation of Support for Principals. The evaluation will provide important information on the implementation and impacts of intensive professional development for principals. The Institute of Education Sciences (IES) at the U.S. Department of Education (ED) has contracted with Mathematica Policy Research and its subcontractors, the American Institutes for Research (AIR), Social Policy Research Associates (SPR), and Pemberton Research to conduct this evaluation.

The evaluation will include implementation and impact analyses. The implementation analysis will draw on information on the scope and sequence of the professional development and data from principal surveys to describe principals' professional development experiences.[1] The impact analysis will be based on a random assignment design in which participating schools in each district are randomly assigned to a treatment group whose principals are offered intensive professional development or to a control group whose principals are not. The impact analysis will draw on data from principal and teacher surveys, principal daily logs, and district administrative records to estimate the impacts of the professional development on principal performance, school climate, teacher retention and performance, and student achievement.

A. Justification

1. **Circumstances necessitating the collection of information**

a. **Statement of need for a rigorous evaluation of principal professional development**

Principals play an important role in the academic performance of students in their schools (Hallinger and Heck 1998; Harris et al. 2010; Knapp et al. 2006; Leithwood et al. 2004), and there is widespread interest in the potential of intensive principal professional development programs to improve principals' performance. However, little is known about the effectiveness of these programs and their ability to improve principals' leadership skills and school quality (Huff et al. 2013). The data collection described in this request will provide essential information for describing implementation of a principal professional development and providing rigorous estimates of its effects on principal performance, school climate, teacher retention and performance, and student achievement.

Legislative authorization for this evaluation is found in Title II, Part A of the Elementary and Secondary Education Act (ESEA), Section 2121-2123, as amended by the No Child Left Behind Act (NCLB) (20 USC 6621-6623). Title II, Part A of ESEA provides funding to states to carry out professional development activities designed to improve the quality of principals. Part F, Section 9601 of ESEA permits program funds to be used to evaluate activities authorized under the act.

---

[1] Data on the scope and sequence of the professional development will be collected from the professional development provider and will not impose burden on study participants.

b. Study design and research questions

To learn about the effectiveness of intensive principal professional development, the study team will evaluate an intensive professional development program using a rigorous random assignment design. In fall 2014 we held a competition to select a promising principal professional development program, and, with input from a panel of experts, selected the Center for Educational Leadership (CEL) at the University of Washington. In the coming months we plan to recruit 100 elementary schools from 10 school districts across the country and randomly assign them to a treatment group whose principals will be offered CEL's principal professional development program or to a control group whose principals will not be offered this program.[2] We will support and monitor the program for two school years to ensure high quality implementation, and will collect comprehensive data on program implementation and school, principal, teacher, and student outcomes. We will use these data to compare outcomes between treatment and control group schools to obtain rigorous evidence on the effectiveness of the program.

The study will address the following research questions about the implementation and impacts of the CEL principal professional development program:

- *Implementation.* What are the professional development experiences of principals in the study?

- *Impacts.* What are the impacts of intensive principal professional development on (a) school climate and educator behaviors, (b) teacher retention and effectiveness, and (c) student achievement and behavior?

c. The CEL principal professional development program

CEL's professional development program is based on a theory of action that principals' instructional-focused leadership boosts teaching quality and raises achievement. According to the program's logic model (Exhibit A.1), if principals have a clear understanding of what quality instruction looks like, as well as a common language to describe and promote elements of quality instruction, and if they know how to lead instructional improvement, instructional practice will become more effective and student achievement will improve.

CEL seeks to develop principals' knowledge of high-quality instruction and to equip them to lead their teachers to deliver quality instruction, by honing their skills in three content areas:

1. Improving teacher instruction through observation, analysis of data (including information from district evaluation systems), and implementation of "inquiry cycles" that generate useful feedback for teachers (*instructional leadership*);

---

[2] To increase the chances that treatment and control groups are well balanced in terms of key baseline characteristics, before random assignment we will organize participating schools in each district into blocks, each of which contains two schools with similar baseline characteristics. Characteristics we will consider in the formation of blocks include (1) grade span; (2) school type (charter, magnet, or traditional public school); (3) average 3rd-, 4th-, and 5th-grade reading and math test scores from the 2013–2014 school year; (4) principal experience; (5) school enrollment; (6) percentage of students eligible for free or reduced-price lunch; and (7) percentage of students who are nonwhite.

2.   Guiding staffing plans and staff development (*human capital management*); and

3.   Creating a school-wide culture of learning to facilitate improved academic success for all students (*organizational leadership*).

Exhibit A.1. Logic model for the CEL Program

Professional development
- Summer institute
- Formal, in-person trainings
- One-on-one coaching
- Online PLCs

Improved leadership practices

Instructional leadership
- Increase the observation and analysis of instruction
- Use evidence to identify areas for growth
- Provide more frequent, better-targeted, and actionable teacher feedback

Human capital management
- Use evidence to identify PD and supports that are better targeted to staff needs
- Use evidence to develop school and staff growth plans

Organizational leadership
- Regularly communicate plans for, and progress toward, school improvement

Teacher changes
- Improved instruction

Schoolwide changes
- Improved/ targeted PD and support for teachers
- Improved school climate and culture

Student learning
- All students, regardless of race, class, language, and disability, achieve at the highest levels

CEL's professional development sessions aim to sharpen principals' skills in observing classrooms and identifying areas in which teachers need to improve. The professional development sessions are tailored to each district's instructional framework and teacher evaluation process, using crosswalks between the different frameworks. CEL also provides individual coaching and professional learning communities (PLCs) aligned with the group sessions.

While the length of CEL's program varies across districts, for the purposes of this study CEL will provide its program for 24 months, beginning with a summer institute in 2015.[3] During the 4-day summer institute, CEL will introduce principals to core program content; establish relationships between principals and coaches; analyze school data; and help principals plan specific leadership actions. After the summer institute, CEL will support principals in treatment schools throughout the 2015–2016 school year, to help them apply lessons from the summer institute in their schools. This support will include 8 one-day professional development sessions in each district; 10 half-day one-on-one coaching sessions (primarily in-person but some may be

---

[3] The institute will be offered on two dates, to help ensure all treatment group principals can attend.

virtual); and quarterly online PLCs. In this way, principals will interact at least twice a month with the professional development provider.

The second year of the program will begin with a summer institute in 2016 for any principals who are new to one of the 50 treatment schools in the study. Principal coaching will continue for all principals in treatment group schools throughout the 2016–2017 school year.

d. Data collection needs

This study includes several data collection efforts that are summarized in Table A.1. The purposes and uses of these data are described in Section A.2 below.

Table A.1. Study data collection activities, by data source

Data source	Mode, timing, and respondent	Key constructs and outcomes
Principal survey (treatment and control groups)		
Annual surveys	30-minute web-based survey, with telephone option and in-person follow-up, administered spring 2016 and spring 2017 to treatment and control group principals	Amount and usefulness of professional development received by principal during the current school year and the preceding summer
		Principals' leadership practices during the school year (human capital management, instructional leadership, and organizational leadership)
		District and school context (resource availability, district policies affecting principal practice)
		Principal background characteristics (demographic traits, educational degrees, certification/licensure, and professional experience)
Principal log (treatment and control groups)		
Daily logs	15-minute web-based logs completed on 5 consecutive days by treatment and control group principals, during four weeklong periods of the 2015–2016 and 2016–2017 school years	Time spent by principal on professional development activities during the school year
		Time spent by principal on various leadership practices (human capital management, instructional leadership, organizational leadership, or other) during the school year
Teacher survey (treatment and control groups)		
Annual surveys	30-minute web-based survey, with telephone and hard-copy options and in-person follow-up, administered to treatment and control group teachers in spring 2016 and spring 2017	Principals' use of leadership practices and quality of principals' instructional and organizational leadership during the school year
		Amount and usefulness of formal and informal professional development received by teachers during the school year
		Amount and usefulness of instructional feedback teachers received from principal or other school leaders during the school year
		Teachers' use of student achievement data and changes to instructional practice (teaching methods, student work, student assessment, and topics or materials) during the school year
		School climate and culture (professional climate, staff collaboration, student engagement and belonging, family engagement, and school safety)
		Teacher background characteristics (demographic traits, educational degrees, certification/licensure, and professional experience)
District records (treatment and control groups)		
Staff records	Electronic grade and school assignment data for prior and current school years, and background characteristics for prior school year, requested from all districts in fall 2015, fall 2016, and fall 2017	Principal and teacher mobility, teacher retention, teacher hiring, and principal and teacher background characteristics
Staff performance records	Electronic performance data for prior school years, requested from all districts in fall 2016 and fall 2017[b]	Teachers' instructional effectiveness and principals' leadership effectiveness

Data source	Mode, timing, and respondent	Key constructs and outcomes
Student records	Electronic student records data for prior school years, requested from all districts in fall 2016 and fall 2017 [b]	Student achievement (test scores), behavioral outcomes (attendance and disciplinary records), and background characteristics (demographic traits, grade level, English language learner status, and special education status)

[b]We will request teacher and principal performance data and student records for the 2014–2015 and 2015–2016 school years in fall 2016 and for the 2016–2017 school year in fall 2017.

e.   Time line for data collection activities

The evaluation is expected to be completed in five years, with 2.5 years of data collection. Table A.2 shows the schedule of data collection activities.

2.   Purposes and uses of the data

To address the study's research questions, the evaluation will collect and analyze data from several sources. IES will use these data to better understand principals' professional development experiences, and to estimate the impact of the professional development on school, teacher, and student outcomes. Table A.3 lists research questions and specific data sources that will be used to answer them—in this table we have broken Research Question 2 into three parts, to correspond to three sets of outcomes we will examine. We describe how the study will use each data source below. Information will be collected by Mathematica Policy Research and its partners AIR, SPR, and Pemberton Research, under contract with ED [contract number ED-IES-14-R-0008].

Table A.2. Time line for data collection activities

Activity	Date
Data for 2015-2016 school year	
Collect principal daily logs	9/2015 through 5/2016
Collect principal and teacher survey data	3/2016 through 5/2016
Collect district records[a]	8/2015 through 12/2015; 8/2016 through 12/2016
Data for 2016-2017 school year	
Collect principal daily logs	9/2016 through 5/2017
Collect principal and teacher survey data	3/2017 through 5/2018
Collect district records[b]	8/2017 through 12/2017

[a]In fall 2015, we will only collect data on teacher assignments for fall 2015. In fall 2016 we will collect district records data for the prior (2014–2015) school year, as well as teacher assignment data for the 2015–2016 school year.

[b]In this data collection, we will also collect district records data on teacher and principal assignments for the 2017–2018 school year.

Table A.3. Research questions and data sources

Research question	Data source(s)
1. Implementation What are the professional development experiences of principals in the study?	Principal surveys
2. Impacts a. *Impacts on intermediate outcomes.* What are the impacts of intensive principal professional development on school climate and educator behaviors?	Principal surveys Principal daily logs Teacher surveys

Research question	Data source(s)
b. *Impacts on teacher outcomes.* What are the impacts of intensive principal professional development on teacher retention and effectiveness?	District records Staff records Staff performance records
c. *Impacts on student outcomes.* What are the impacts of intensive principal professional development on student achievement and behavior?	District records Student records

Principal Surveys. Surveys administered to treatment- and control-group principals in years 1 and 2 will collect information on their leadership practices and perceptions of school climate. We will use this information to examine the effect of professional development on leadership practices and school climate (research question 2a). We will also ask principals to report the amount, content, and usefulness of their professional development experiences. We will use this information to document treatment and control group principals' perceptions of their professional development experiences (research question 1). In addition, principals will provide data on their backgrounds, professional experience and educational attainment, and the district context in which they operate (including, for example, barriers to school improvement). We will use this information to document the context of principals' professional development experiences (research question 1).

Principal Daily Logs. Daily logs will collect information on how principals allocate their time across different categories of leadership practice. Logs will be administered to treatment and control groups on five consecutive days during four weeks of school years 1 and 2. We will use this information to examine the effect of professional development on daily time-use in different domains of leadership (research question 2a). As compared to annual surveys, daily logs have been shown to more accurately capture frequently occurring behaviors, such as day-to-day instructional leadership and managerial responsibilities, because they allow for measurement of daily fluctuations and are less susceptible to recall error (Camburn et al. 2010a, 2010b). Principal logs will complement annual principal surveys by ensuring that the study team can examine impacts on how principals are actually spending their time on leadership behaviors on a given day, rather than relying only on annual recall; and allowing detection of changes in principals' leadership behaviors within a year.

Teacher Surveys. Surveys administered to teachers in years 1 and 2 will collect information on teachers' experiences during the school year, including the amount and perceived usefulness of professional development and instructional support they received, instructional improvement and data-use practices, and perceptions of their principal's leadership practices and school improvement efforts. We will use this information to examine the effect of principal professional development on teacher practices and teacher perceptions of principals' leadership practices. Teachers will also be asked about their schools' working conditions, staff collaboration and support, and academic culture during the school year. We will use this information to examine the effect of principal professional development on school climate (research question 2a). Finally, teachers will provide data on their backgrounds, educational attainment, and professional experience. We will use this information to provide context for understanding principal's professional development experiences (research question 1) and to examine how principal professional development affects the mix of teachers hired and retained in the study schools (research question 2b).

District Records. We will collect data on principals, teachers, and students from the district records listed below. We will use this information to estimate effects of principal professional development on staff mobility, retention, and performance (research question 2b) and on student achievement and behavior (research question 2c), and we will also use district data to draw a sample of teachers to take the teacher survey in years 1 and 2.

- **Staff records.** We will collect data on principals' school assignments and teachers' school assignments and the grades and subjects taught. We will use this information to examine effects of principal professional development on principal turnover and teacher retention (research question 2b). We will also use it to draw a random sample of teachers to take the survey in years 1 and 2 of the study. We will collect available information on characteristics of the study sample, such as principal and teacher demographics (age, sex, race, and ethnicity), educational attainment (certifications, degrees, and scores on licensure or certification exams), and years of teaching experience. We will use this information to describe the study context in the implementation analysis (research question 1), and to examine whether impacts of principal professional development vary depending on baseline values of these characteristics, such as principal experience (research question 2).

- **Staff performance records.** We will collect principal and teacher performance ratings from district evaluation systems. These may include composite ratings, observation ratings, teacher value-added scores, school value-added scores, or scores on student learning objectives. We will use this information to examine the effect of principal professional development on teacher and principal effectiveness as assessed by their district systems (research question 2b). In conjunction with staff records described above, we will also use this information to examine effects on hiring and retention of effective teachers, dismissal of ineffective teachers, and ways that principals assign teachers of different performance levels to classes (research question 2b).

- **Student records.** We will collect information from student records for years 1 and 2, as well as the school year before the program started. We will obtain student standardized test scores in math and English/language arts for all tested grades, which we will use to examine the effect of principal professional development on student achievement (research question 2c). We will also obtain data on student attendance and discipline, which we will use to examine effects on student behavior. Finally, we will collect data on student demographics (age, sex, race, and ethnicity) and participation in other programs (free or reduced-price lunch eligibility, English Language Learner status, and individualized education plan status). We will use this information to describe students in the study sample and use baseline values of these measures and student test scores to improve precision of estimates in the impact analysis (research question 2).

Data collection in year 1. We will use information collected during year 1 to examine principals' professional development experiences (research question 1) and impacts of principal professional development during the initial year of program implementation (research question 2). We will use information on year 1 educator behaviors and teacher and student outcomes to inform research question 2, by examining whether principal professional development resulted in differences between the treatment and control groups after the first year of the intervention's implementation.

Data collection in year 2. We will use the information collected during year 2 to examine the impacts of intensive principal professional development one year later, after two years of the intervention's implementation (research question 2). More specifically, we will use the data collected in year 2 to examine whether principal professional development resulted in differences in school climate, teacher and principal practices, and teacher, principal, and student outcomes after two years of professional development.

B. Collection of information requiring statistical methods

1. Respondent universe and sampling methods

The respondent universe for the schools and districts in the study includes high-poverty elementary schools across the country and the districts in which they are located, while the respondent universe for the principal, teacher, and student samples includes all principals, teachers, and students in the schools that participate in the study. The study team will select the sample in five stages: (1) We will identify potential districts and elementary schools to include in the study, and reach out to these districts and schools to request their participation, with the goal of recruiting a sample of 100 high-poverty elementary schools in 10 districts. (2) We will randomly assign participating schools in each district to a treatment group whose principals will be offered CEL's principal professional development program in the 2015–2016 and 2016–2017 school years or to a control group whose principals will not. (3) We will collect survey data and daily logs from all principals in the 100 schools participating in the study, and will collect administrative data on these same principals from districts. (4) We will collect administrative data on all teachers in the study schools, and draw a stratified random sample of 12 teachers in each school to complete the teacher survey. (5) We will collect administrative data on all students in the study schools. Below we describe each stage in greater detail.

a. Selection of districts and schools

The study will include a purposive sample of high-poverty elementary schools, in districts that are not already offering intensive principal professional development. We will identify potential districts and elementary schools to include in the study, and reach out to these districts and schools to request their participation, with the goal of recruiting a sample of 100 high-poverty elementary schools in 10 districts. To increase the chance that at least 10 schools in each district will be eligible and willing to participate after further screening, we plan to limit the list of eligible districts to those with at least 20 high-poverty elementary schools that are not already offering intensive professional development, defined as follows:

- **High-poverty.** We defined schools as high-poverty if they had at least 40 percent of students eligible for free or reduced-price lunch, based on statistics from the 2011–2012 Common Core of Data (the most recent data available). We used the 40 percent threshold because it is the same threshold used to determine school wide Title I status.

- **Elementary schools.** We defined schools as elementary schools if they had a starting grade between prekindergarten and 4 and an ending grade between 4 and 6. This definition ensures that study schools will all include at least one or two tested grades. It also excludes schools with middle and high school grades (beyond grade 6), due to concerns that the effects of the professional development could be different at the secondary level.

- **Not offering intensive professional development.** We excluded districts that we knew were already implementing intensive principal professional development or other major leadership initiatives district-wide, including the National Institute for School Leadership (NISL), the Wallace Foundation's Principal Pipeline Initiative, the Wallace Foundation's Principal Supervisor Initiative, the Bill & Melinda Gates Foundation's Intensive Partnership Program, and the University of Virginia School Turnaround Program. We will further screen the districts that we contact for recruiting to assess whether they are implementing intensive principal professional development that we are not yet aware of.

Applying the above criteria resulted in a list of 122 eligible districts. To decide which of the remaining districts to target for recruitment, and to ensure a geographically diverse sample with a mix of large and mid-sized districts, we first divided eligible districts into eight bins, based on district size (midsize and large) and geographic region (Northeast, Midwest, South, West). We designated districts as midsize if they had 20 to 40 elementary schools and large if they had more than 40—this threshold roughly evenly divided all eligible districts in the sample between the two groups. We then randomly ordered the eligible districts from each bin, and selected districts to target for initial recruiting in rough proportion to the number of eligible districts in each bin, resulting in a list of 60 districts to target for initial recruitment, with the goal of obtaining a final sample of 10 districts. Within each participating districts, we will work with district staff to identify a set of approximately 10 high-poverty elementary schools that are willing and able to participate in the study.

b. Random assignment of schools

After obtaining each district's formal agreement to participate in the study and identifying the schools that will participate, we will randomly assign these schools to a treatment group whose principals are offered the CEL professional development or a control group whose principals are not. We anticipate that 50 schools will be assigned to the treatment group and 50 to the control group.

c. Principal sample

The sample of principals will include all 100 principals in study schools at the time of random assignment, as well as any replacement principals that join the study schools during the two years of the study. Given expected annual principal turnover rates of 27 percent (Goldring et al. 2014), we expect approximately 27 new principals to join the study sample, resulting in a total sample of 127 principals.

Principal sample for principal survey, daily logs, and district records collection. To estimate program impacts, we will collect data from all original and replacement principals in treatment and control group schools. This data collection will include annual principal surveys; principal daily logs administered on five consecutive days, during four weeklong periods of the

school year; and district records on principals' backgrounds, mobility and performance. The respondent universe for this data collection will include all of the approximately 100 original and 27 replacement principals from the 100 schools in the study. We expect that approximately 85 percent of principals in the sample will complete the survey and the daily logs, and we expect to obtain administrative records data for all principals in the sample.

d.   Teacher sample

The respondent universe for data collection from teachers will include all of the approximately 3,000 teachers in treatment and control group schools during the two years of the study (roughly 30 teachers per school).

Teacher sample for teacher survey. To document teacher perceptions of principal leadership, teacher professional development and instructional improvement, and school-wide climate and culture, we will administer a 30-minute survey to a random sample of 12 teachers per school. To select the sample, we will draw a random sample, stratified by grade level, of 12 teachers per school to complete the survey in each study year. We anticipate the sample for each of the two surveys will include 1,200 teachers across the 100 participating schools. We expect that approximately 85 percent of teachers in the survey sample will complete the survey.

Teacher sample for administrative records collection. We will collect district administrative records on teachers' backgrounds, mobility, placement, and performance for all of the approximately 3,000 teachers in study schools during the two years of the study. We expect to obtain administrative records data for all teachers in the sample.

e.   Student sample

The respondent universe for student data collection will include all students in study schools during the two years of the study. We anticipate this universe will include approximately 50,000 students (about 500 students per school). To examine effects on student achievement, we will collect student standardized test scores in math and English/language arts for all tested grades (3 through 6)—we anticipate this universe will include approximately 25,000 students. To examine effects on student behavior, we will obtain attendance and discipline records. We will also collect data on student demographics (age, sex, race/ethnicity, and eligibility for free or reduced-price lunch) and educational statuses (grade level, English Language Learner status, and individualized education plan status) to describe the students in the study sample and improve the precision of the impact estimates. We expect to obtain student records data for 99 percent of students in the sample.

2.   Procedures for the collection of information

a.   Statistical methodology for stratification and sample selection

As described in our response in section B.1, we will select a purposive sample of 10 districts and 100 high-poverty elementary schools that are not already implementing intensive principal professional development and that are willing and able to participate in the evaluation. We will then randomly assign half of the participating schools in each district to a treatment group whose principals are offered intensive principal professional development and half to a control group that are not. We will not draw a random sample of districts or schools for the study, which is

infeasible because districts and schools must be willing to participate and implement a principal professional development program. Below we describe our selection of the sample of principals, teachers, and students for each data collection activity.

Selection of principal sample. We will not randomly sample principals for any data collection activities. Instead, we will attempt to collect survey data, daily logs, and district records from all principals in both treatment and control group schools.

Selection of teacher sample. We will draw a random sample of teachers for the teacher survey, stratified by grade to ensure representation of teachers from all grade levels. We will attempt to collect administrative records data on all teachers in both treatment and control group schools.

Selection of student sample. We will not randomly sample students for the collection of district administrative records. Instead, we will attempt to collect administrative records data on all students in both treatment and control group schools.

b.   Estimation procedures

The study will include an implementation analysis that details the professional development experiences of participating principals and an impact analysis that examines the impacts of the principal professional development on school climate, educator behavior, teacher performance and retention, and student achievement.

Implementation analysis

The implementation analysis will contextualize findings from the impact analysis, detailing the fidelity of program implementation and challenges encountered, identifying mechanisms through which the professional development might improve outcomes, and documenting the extent of the contrast between the intervention and the counterfactual (that is, whether the experiences of principals in the treatment group differed significantly from those of principals in the control group).

Descriptive analysis of professional development. Drawing on information on the intended scope and sequence of the study's professional development gathered by the study's technical assistance team, we will first describe the intervention as planned, including its objectives, and the timing, hours of professional development (session-based and coaching), delivery mode, and content of support.[4] We will draw on principal surveys to describe the professional development received by treatment group principals. We will describe the number of hours and content of professional development received and principals' perceptions of the quality or usefulness of their professional development experiences. Using information from the principal surveys, teacher surveys, and district records data, we will also describe the characteristics of the districts in which the intervention was delivered.

Comparison of the experiences of treatment and control group principals. Using data from the year 1 principal survey, we will describe the professional development received by the

---

[4] These data will be collected from the professional development provider by the study's technical assistance team and will not impose burden on study participants.

treatment group principals relative to that received by control group principals. The analysis will detail the content and frequency of professional development both groups received and test for statistically significant differences between the two groups.

Impact analysis

The impact analysis will examine the effects of CEL on the principal, teacher, and student outcomes in Table B.1. Key outcomes of interest include (1) teachers' receipt of professional development and perceived quality and frequency of feedback from the principal; (2) school climate outcomes, including staff collaboration and perceptions of support; (3) teachers' perceptions of principals' leadership practices, school climate, and educator behaviors; (4) teacher effectiveness (as measured by the district's central office) and teacher retention; and (5) student test scores, attendance, and behavior.

Table B.1. Principal, teacher, and student outcomes for the impact analysis

Level	Outcomes
Principal	• District central office performance ratings

Teacher	• Perceptions of school climate
	  – Staff collaboration
	  – Perceived support
	• Perceptions of principals' leadership practices
	• Teachers' receipt of professional development and principal feedback (quality and usefulness)
	• Retention
	• Performance ratings
	  – All current teachers
	  – New teachers
	• Assignment
	  – Difference between teacher performance in high-stakes, low-stakes grades
	  – Variance of average teacher performance across grades (to examine whether high- and low-performing teachers grouped within grades)

Student	• Achievement • Attendance • Behavior

We will estimate impacts using the following model:

$$Y_{ij} = \alpha + \beta T_j + \delta P_{ij} + \gamma Z_j + \varepsilon_{ij} \quad (1) \quad ,$$

where Y_{ij} is the outcome of interest for individual (principal, teacher, or student) i in school j; α is an intercept term; T_j is an indicator equal to one if the school is assigned to the treatment group and zero otherwise; P_{ij} is a vector of baseline school- or student-level characteristics, and Z_j is a vector of fixed effects corresponding to the study's random assignment blocks; δ and γ are coefficient vectors; and ε_{ij} is a random error term, clustered at the school level. The coefficient β represents the average impact of the principal professional development.[5] We will apply a weighting scheme that gives an equal weight to each of the schools in the sample (regardless of

---

[5] We will not adjust for principal and teacher characteristics in this model. Because the intervention could influence the composition of principals and teachers (through retention or mobility), observed characteristics of principals and teachers could reflect program effects, and including them could lead to an incorrect estimate of program impacts.

the number of principals, teachers, or students in the sample in each school).Thus, impact estimates will reflect the effect of the professional development for the average school in the sample, rather than for the average teacher or student.

For teacher assignment, we will investigate how assignment to the CEL program affects the way principals assign teachers to classes—for instance, whether it leads principals to pair higher-performing teachers with lower-performing teachers in the same grade (to facilitate mentoring), or whether it leads them to reassign low-performing teachers to grades not covered by high-stakes tests. We will also examine principal and teacher performance ratings as measures of effectiveness, but we will interpret these data with caution. Some districts will rate principals at least partly based on school value added. For new principals, this rating is affected by the teachers inherited from prior principals (Chiang et al. 2012). Similarly, teacher performance ratings may have two limitations: (1) our experience suggests that nearly all teachers are given 'satisfactory' ratings, so there may be little variation in the data; and (2) teacher evaluations typically incorporate ratings by principals, so differences in ratings could reflect effects of the professional development program.

Our main impact estimates will reflect the impact of the professional development on the schools whose principals were offered the professional development, whether or not they actually participated for the full two years. To examine the effects on schools that received the full intended "dosage"—that is, treatment schools whose principals received both years of the CEL professional development program—we will estimate local average treatment effects (Imbens and Angrist 1994), using treatment status as an instrumental variable for the proportion of total hours of intended CEL activities the principal attended. Even schools with principal turnover during the study could receive the full intended dosage if both the original and replacement principals attend all offered CEL activities during their time at the school.

Subgroup analyses. Because new principals may be more in need of and more receptive to coaching and formal group professional development sessions than more experienced principals, we will conduct subgroup analyses separately examining the effects of the professional development for novice and experienced principals. Statistical power for subgroup analyses will be lower than for the full sample—subgroup findings will thus be less definitive than those for the full sample, but can provide suggestive evidence of impacts for these groups.

c.   Degree of accuracy needed

We estimate that the target sample of 100 schools will yield a minimum detectable impact on student test scores as small as 0.10 standard deviations for the full sample of students in tested grades (Table B.2). Similarly, we estimate a minimum detectable effect of 0.15 standard deviations for a subsample of 50 schools—for example, when estimating impacts separately for districts with high and low levels of principal experience. A sample size of 100 schools will also enable us to detect reasonably sized impacts on key principal and teacher outcomes—for example, we estimate we will be able to estimate an impact of 8 percentage points on the percent of time principals spent on instruction-related tasks (about a 40 percent increase from the expected percent of time spent by control group teachers) and an impact of 11 percentage points on teacher retention (about a 20 percent increase from the expected retention rate of control group teachers). Under our plans to survey a sample of 12 teachers per school (described in more

detail in Section B.1 above), we will achieve a minimum detectable impact of 9 percentage points on the percent of teachers who strongly agree that their principals give useful feedback on teaching at least once a month (less than half the expected 19 percentage point difference in this outcome between schools with medium-rated and highly-rated instructional leaders [Quint et al. 2007]).

Table B.2. Minimum detectable impacts on student, teacher, and principal outcomes, sample of 100 schools

Data source	Outcome	Sample size	Minimum detectable impact	
			Full sample	Subgroup (50% sample)
District records	Student test scores	25,000 students	0.10 SDs	0.15 SDs
District records	Teacher retention	3,000 teachers	11 pct. points	16 pct. points
Teacher survey (sample of 12 teachers per school	% of teachers who strongly agree that their principals give useful feedback on teaching	1,200 teachers	9 pct. points	14 pct. points
Principal survey	% of time spent on instructional leadership	100 principals	8 pct. points	12 pct. points

Notes: Calculations assume (1) 80 percent power and a 5 percent significance level for a two-tailed test; (2) schools will contain an average of 250 students in tested grades (3 through 6); (3) schools will contain an average of 30 teachers; (4) we will obtain outcome test score data for 99 percent of students in tested grades and retention data on all teachers, and 85 percent of principals and teachers will respond to the survey; (5) the intracluster correlation is 0.16 for student and teacher outcomes; (6) covariates explain 80 percent of the between-school variance and 50 percent of the within-school variance of student test scores, 20 percent of the between-school variance and 10 percent of the within-school variance for teacher outcomes, and 20 percent of the between-school variance for principal outcomes; and (7) 60 percent of teachers in the control group will be retained each year (Glazerman et al. 2013), the mean and standard deviation of the percent of time principals spend on instruction-related tasks will be 19 percent and 15 percentage points, respectively (May et al. 2012), and 16 percent of control group teachers will report that they strongly agree that their principals give useful feedback on teaching at least once a month (Quint et al. 2007). Assumptions on the clustering of outcomes and the explanatory power of covariates for the student analysis are based on data from five large random assignment evaluations in K–12 education (Deke et al. 2010). School size assumptions are based on tabulations for high-poverty elementary schools from the Common Core of Data.

SDs = standard deviations.

d.   Unusual problems requiring specialized sampling procedures

We do not anticipate any unusual problems that will require specialized sampling procedures.

e.   Use of periodic (less frequent than annually) data collection cycles to reduce burden

Because the effects of intensive principal professional development may change over time as principals apply the lessons of the program to their leadership practices, we plan to measure the intermediate effects of the program after one year of implementation and the longer term effects after two years of implementation. To fully capture both short- and longer-term effects, we will need to collect all data two times, once for each study year.

f.   Who will collect the data and how it will be done

Following is a description of how the study team will collect data on treatment and control groups from four specific sources: (1) principal surveys, (2) principal daily logs, (3) teacher surveys, and (4) district records.

Principal surveys. The data collection team will administer a 30-minute web-based survey to treatment and control group principals. We will administer principal surveys in spring 2016 (the end of year 1) and spring 2017 (the end of year 2). We will supplement the initial survey with telephone and in-person follow-up. Principals will respond to close-ended prompts about their professional development experiences during the school year and preceding summer, their leadership practices and school safety during the school year, the district and school contexts in which they operated during the school year (for example, the availability of resources), and their social backgrounds (demographic traits, professional experience, and educational attainment).

Principal daily logs. The data collection team will administer a 15-minute web-based log to treatment and control group principals during the 2015-2016 and 2016-2017 school years. On five consecutive days, during four weeklong periods of each school year, we will ask principals to report how they allocated their time during the day in fifteen-minute increments across different categories of leadership practice. This will yield daily principal time-use data for a total of 20 days per principal per school year.

Teacher surveys. The data collection team will administer a 30-minute web-based survey to a random sample of 12 teachers per school, with telephone and hard-copy options. We will administer teacher surveys in spring 2016 and spring 2017. Teachers will respond to close-ended prompts about their experiences during the school year, including the professional development and instructional support they received, practices they used and instructional improvements they made, principals' leadership practices, school-wide climate and culture, and their social backgrounds (demographic traits, professional experience, and educational attainment).

District records. The data collection team will collect data on principals, teachers, and students from the school districts of treatment and control group principals, from the following extant administrative records:

- **Staff records.** In fall 2015, fall 2016, and fall 2017, the data collection team will obtain district records of teachers' and principals' school assignments and the grades and subjects taught by teachers in the prior school year and in the current school year (as of the time of data collection, for purposes of drawing the sample for the teacher survey). We will also collect available records of principal and teacher demographics (age, sex, race, and ethnicity), educational attainment (certifications, degrees, and scores on licensure or certification exams), and years of teaching experience for the prior school year. We will collect district records for all teachers and principals in treatment and control group schools.

- **Staff performance records.** In fall 2016, the data collection team will collect principal and teacher performance ratings for the 2014-2015 and 2015-2016 school years, and in fall 2017, they will collect these data for the 2016-2017 school year. The data may include composite ratings, observation ratings, teacher value-added scores, school value-added scores, and

scores on student learning objectives. We will collect performance records for all teachers and principals in treatment and control group schools.

- **Student records.** In fall 2016 the data collection team will obtain student records data for the 2014-2015 and 2015-2016 school years, and in fall 2017 they will obtain these records for the 2016-2017 school year. These data will include student standardized test scores in math and English/language arts for all tested grades, student attendance, student disciplinary records, student demographics (age, sex, race/ethnicity, and eligibility for free or reduced-price lunch), and student educational statuses (grade level, English Language Learner status, and individualized education plan status). We will collect district records for all students in treatment and control group schools.

3.   Methods to maximize response rates and deal with nonresponse

Below we describe our methods for maximizing response rates and minimizing nonresponse in our collection of extant data from districts and our collection of primary data from principals and teachers.

Extant data collection. To reduce districts' burden in the submission of extant data and maximize response rates, we will allow them to submit data in the most convenient format. Federal rules permit ED and its designated agents to collect student demographic and existing achievement data from schools and districts without prior parental or student consent (Family Educational and Rights and Privacy Act [20 U.S.C. 1232g; 34 CFR Part 99]). To further maximize the response rate and minimize burden on schools, we will follow these federal rules.

Primary data collection. We will employ a multifaceted approach to reducing nonresponse for the study's primary data collection, which includes the principal surveys, principal daily logs, and teacher surveys. Following is a description of our approach, which is also summarized in the box below. To ensure the targeted overall response rates of 85 percent, and to reduce disparities in response rates between the treatment and comparison group samples, we will make these available in multiple modes so that sample members can complete it in the mode that is most convenient for them. We will invite respondents to participate in the survey or log data collections through email messages and mailed letters and conduct non-response follow-up through email messages, telephone calls, mailed letters, and for some districts in-person visits. We will thoroughly test the instruments for such features as clarity, accuracy, length, flow, and wording. We will minimize the length of instruments to what we believe is sufficient to gather key information while not overburdening respondents. To reduce item nonresponse, the web-based questionnaires will include programmed checks alerting respondents to all out-of-range or inconsistent responses they enter. The edit check will give respondents the choice of changing their response, based on guidance provided on the pop-up screen, or leave their answer and continue on to the next question. To further minimize burden on teachers, we will administer teacher surveys to a random sample of 12 teachers per school. We are also proposing burden payments for the principal survey, the principal log, and the teacher survey to partially offset respondents' time and effort in completing them. The payments are proposed because high response rates are needed to make the survey findings reliable, and we are aware that principals and teachers are the target of numerous requests to complete surveys on a wide variety of topics from state and district offices, independent researchers, and ED.

Minimize burden on principals and teachers

- Principals and teachers will be able to access instruments by clicking a web link provided in an email.
- The principal and teacher surveys will take approximately 30 minutes to complete, while each principal log will take about 15 minutes to complete.
- Automated skip patterns will be used to ask questions that are appropriately tailored to each respondent, based on their responses to earlier questions.

Maximize principal and teacher response rates

- We will provide respondents with monetary incentives via mail upon completion of each survey and after completion of each of the four 5-day log data collections.
- We will use a tracking system to monitor response rates during the data collection period.
- We will remind respondents to participate through email messages, phone calls, mailed letters, and in some districts in-person school visits.
- We will provide principals and teachers multiple options for how to complete surveys (web-based and hard-copy).

Principal survey. We will use the following specific procedures to maximize response rates on the principal survey. Principals will access the survey by clicking on a web link provided in an email. At the time of study enrollment, we will collect email addresses and ensure that principals have web access. We will also collect principals' personal telephone numbers during study enrollment, as we have faced challenges directly reaching principals when calling schools on past studies. Follow-up attempts will be made by email and then by telephone. During the telephone follow-up, principals will also be offered the option of providing the data over the phone. For this effort, experienced interviewers will be recruited and extensively trained on data collection procedures, including methods for promoting principal cooperation. For any remaining non-responders, we will work with district coordinators to schedule in-person visits to schools during the last four weeks of principal survey data collection. Experienced staff will visit schools to motivate principals to participate and to provide a flyer listing a web link and incentive information. During each round of data collection, we propose to offer $30 to principals who complete the principal survey. By using these methods, we expect that 85 percent of principals will submit a principal survey per year, based on the outcome of similar prior data collection efforts (Camburn et al. 2010b).

Principal daily logs. We will use the following specific procedures to maximize response rates on the principal logs. Principals will access the web-based log via a link provided in an email. We will offer a webinar-based training (on several dates/times) to demonstrate the daily logs and answer questions. Principals will receive email reminders on logging days that also summarize the number of logs already completed for the week and prompt them to submit any uncompleted logs. During each round of data collection, we propose to offer $10 to principals for each day a log is completed (for a possible total of $50 per round and a total of $200 across all four rounds of data collection in each school year). By using these methods, we expect that 85 percent of principals will submit at least one week's worth of daily log data and 70 percent will submit complete daily log data for all 20 days per year, based on the outcome of similar prior data collection efforts (Camburn et al. 2010a; May et al. 2012).

Teacher surveys. We will use the following specific procedures to maximize response rates on the teacher surveys. Teachers will receive a link to the web-based survey by email. In previous studies in similar settings, we have found that some teachers do not check school email accounts frequently; therefore, teachers will also be given the option of completing a hard-copy

survey, which will be mailed to them at their schools. Over a 12-week data collection period, teachers will receive email and mail reminders. During each round of data collection, we propose to offer $30 to teachers who complete the teacher survey. We will also coordinate in-person school visits with our district coordinator during the last four weeks of data collection to provide teachers with a hard-copy version of the teacher survey (during school faculty meetings, if permitted). This in-person connection has helped motivate teachers to participate in past surveys. By using these methods, we expect that 85 percent of sampled teachers will submit a teacher survey per year, based on the outcome of similar prior data collection efforts (Supovitz et al. 2010).

4. Pilot testing

Principal Survey and Principal Daily Log. We are in the process of conducting pilot testing of the principal survey and principal daily log. We are conducting the pilot tests of the principal survey and principal daily log in two rounds so that any changes from the first round of pilot tests with four principals can be included in the instrument for the second round of pilot tests (with no more than 5 principals). Each principal (no more than nine total) will receive a survey and one daily log to complete.

Teacher Survey. Cognitive pretesting with eight teachers, across two rounds, four in the first and four in the second, has been conducted for the teacher survey. The cognitive pretest sample included teachers from schools districts in three states. It took an average of 28 minutes to complete the teacher survey in its entirety. Survey questions have been revised to improve clarity based on the results of these tests.

5. Individuals consulted on statistical aspects of the design

The study team and Technical Working Group members, listed in Table B.3 were consulted on various aspects of the statistical design:

Table B.3. Individuals Consulted on Study Design

Name	Title and affiliation
Susanne James-Burdumy	Senior Fellow, Mathematica
Melissa Clark	Senior Researcher, Mathematica
Christina Tuttle	Senior Researcher, Mathematica
Mark Dynarski	Founder, Pemberton Research
Eric Camburn	Associate Professor and Senior Researcher, Consortium for Policy Research in Education (CPRE), University of Wisconsin-Madison
Roger Goddard	Professor and Fawcett Chair, Department of Educational Studies, the Ohio State University
Jason Grissom	Assistant Professor and Faculty Affiliate, Center for the Study of Democratic Institutions, Vanderbilt University
Jason Huff	Educational Consultant, Seattle Public Schools
Carolyn Kelley	Professor, Educational Leadership and Policy Analysis, University of Wisconsin-Madison
Jim Kemple	Executive Director of the Research Alliance for New York City Schools; Research Professor at the Steinhardt School of Culture, Education, and Human Development, New York University
Susanna Loeb	Barnett Family Professor of Education, Stanford University
John Nunnery	Executive Director, Center for Educational Partnerships, Old Dominion University
Jeff Smith	Professor, Economics, University of Michigan

Table B.4 lists the people responsible for data collection and analysis:

Table B.4. Data collection and analysis group members

Name	Title and affiliation	Telephone
Susanne James-Burdumy	Senior Fellow, Mathematica	609-275-2248
Christina Tuttle	Senior Researcher, Mathematica	202-238-3323
Melissa Clark	Senior Researcher, Mathematica	609-750-3193
Ben Martinez	Senior Research Scientist, AIR	
Tim Silva	Associate Director of Research, Mathematica	202-484-5267
Pat Nemeth	Senior Survey Researcher, Mathematica	609-275-2294
Tim Novak	Director of Human Services Research Systems, Mathematica	609-936-3257

REFERENCES

Camburn, E., J. Spillane, and J. Sebastian. "Assessing the Utility of a Daily Log for Measuring Principal Leadership Practice." *Educational Administration Quarterly,* vol. 46, 2010a, pp. 707–737.

Camburn, E., J. Huff, E. Goldring, and H. May. "Assessing the Validity of an Annual Survey for Measuring Principal Leadership Practice." *Elementary School Journal,* vol. 111, no. 2, 2010b, pp. 314–335.

Camburn, E. Goldring, J. Sebastian, H. May, and J. Huff. "Conducting Randomized Experiments with Principals: A Case Study." *Educational Administration Quarterly,* forthcoming.

May, H., J. Huff, and E. Goldring. "A Longitudinal Study of Principals' Activities and Student Performance." *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice,* vol. 23, no. 4, 2012, pp. 417–439.

Supovitz, J., P. Sirinides, and H. May. "How Principals and Peers Influence Teaching and Learning." *Educational Administration Quarterly,* vol. 46, no. 1, 2010, pp. 31–56.

Improving public well-being by conducting high quality, objective research and data collection

MATHEMATICA
Policy Research