




Motivating Leadership Change and Improvement: How Principal Evaluation Addresses Intrinsic and Extrinsic Sources of Motivation

Educational Administration Quarterly
2023, Vol. 59(5) 965–1001
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0013161X231188706
journals.sagepub.com/home/eaq



Madeline Mavrogordato¹ ,
Peter Youngs²,
Morgaen L. Donaldson³ ,
Hana Kang⁴ ,
and Shaun M. Dougherty⁵

Abstract

Purpose: This mixed-methods study examined the association between the degree to which principal evaluation systems include intrinsic and extrinsic sources of motivation and principals' perceptions of whether their district's evaluation system promotes leadership change and improvement. We also investigated how principals experience intrinsic sources of motivation in the context of principal evaluation. **Research Methods/Approach:** For

¹Michigan State University, East Lansing, MI, USA

²University of Virginia, Charlottesville, VA, USA

³University of Connecticut, Storrs, CT, USA

⁴University of Colorado Colorado Springs, Colorado Springs, CO, USA

⁵Boston College, Chestnut Hill, MA, USA

Corresponding Author:

Madeline Mavrogordato, College of Education, Michigan State University, 620 Farm Lane, East Lansing, MI 48824, USA.

Email: mavro@msu.edu

our quantitative analysis, we administered surveys to 82 elementary and middle school principals in 21 districts in Connecticut, Michigan, and Tennessee. We used multiple regression analysis to identify factors associated with principals' perceptions of their district evaluation systems. For our qualitative analysis, we selected six principals from the 82 survey participants. We used the constant comparative method and a consensus approach to coding the interview data; this enabled us to identify linkages between key codes and broader themes. **Findings:** Our quantitative analyses indicated that principals' perceptions of whether their evaluation system promoted leadership improvement were strongly associated with the degree to which they reported that their evaluation system included intrinsic sources of motivation. Our qualitative analysis revealed clear differences among principals with regard to their interactions with district administrators, the nature of their principal professional development activities, and their experiences with autonomy, feedback, and district expectations. **Implications:** Principal evaluation and professional development are under-researched topics. This study identified mechanisms through which principals seem to find evaluation intrinsically motivating. This is one of the first studies to empirically test Firestone's argument that evaluation systems that support intrinsic motivation are more effective than those that emphasize extrinsic motivation.

Keywords

principal evaluation, intrinsic motivation, extrinsic motivation, mixed methods, professional development

In the past decade, interest in principal evaluation has increased substantially in the United States (U.S.). In response to the 2008 federal Race to the Top initiative (RTTT), almost every state approved legislation that altered how principals are evaluated and supported (Donaldson et al., 2021a; Fuller et al., 2015; Jacques et al., 2012) and the Wallace Foundation launched a major initiative to overhaul principal supervision in six large school districts (Goldring et al., 2020). In response, some researchers have used state-level data to examine how principals' evaluation ratings are associated with their schools' characteristics, student demographics, and other measures of principal performance (Grissom et al., 2018; Henry & Viano, 2015; Hermann & Ross, 2016; McCullough et al., 2016). Other scholars have investigated changes in the role of principal supervisors at the district level (Anderson & Turnbull, 2016; Kimball et al., 2015;

Rubin et al., 2021), and some studies have considered associations between principal evaluation and leadership practice (Goldring et al., 2020; Hamilton et al., 2012; Sun & Youngs, 2009).

At the same time, few studies have captured principals' perspectives regarding how they are experiencing evaluation under the new policies implemented in the wake of RTTT. In particular, there has been little research on whether school leaders perceive principal evaluation systems as promoting leadership change and improvement; that is, whether principals view these systems as fair, aligned with their own vision of effective leadership, and prompting them to change how they allocate time to various leadership activities. In addition, there has been little research on how principal evaluation systems are designed to motivate principals to improve, and how these sources of motivation shape principals' perceptions of whether their evaluation system enhances leadership change and improvement.

Further, we know little about whether principals vary in the degree to which they perceive that their districts' principal evaluation systems include intrinsic sources of motivation, such as goal setting, feedback, and professional development linked to principal evaluation results. In this paper, we report on a mixed methods study of elementary and middle school principals in Connecticut, Michigan, and Tennessee that was designed to address these shortcomings in the research literature. We define intrinsic motivation as the act of engaging in behavior because one derives enjoyment or satisfaction from it and, drawing on others' work, argue that goal setting and feedback are key factors that can promote intrinsic motivation (Deci et al., 1996; Firestone, 2014). We define extrinsic motivation as the act of engaging in behavior because of an outside expectation, pressure, or incentive to do so (Donaldson, 2020; Firestone, 2014).

Our mixed methods study was designed to address the following questions:

1. What is the association between the degree to which principals report that their districts' principal evaluation systems include intrinsic sources of motivation and extrinsic sources of motivation and their perceptions of whether their district's evaluation systems promote leadership change and improvement?
2. How do principals describe and experience intrinsic sources of motivation in the context of their principal evaluation system?

In this paper, we used surveys and interviews with a sample of principals to probe motivation and improvement in the context of principal evaluation. To answer the first question, we drew on survey data from elementary and middle school principals in 21 districts across Connecticut, Michigan, and Tennessee

to examine associations between their perceptions of whether and how principal evaluation included intrinsic sources of motivation (i.e., goal setting, feedback, and professional development) and/or extrinsic sources of motivation (i.e., principal accountability, personnel decisions); and whether principal evaluation enhanced leadership change and improvement. Based on our quantitative results, we utilized interview data from a subset of six survey participants to qualitatively explore how school leaders experienced intrinsic sources of motivation in their principal evaluation systems.

Principal Evaluation, Professional Development, and Leadership Practice

Several researchers examined principals' perceptions of and experiences with principal evaluation prior to the 2008 RTTT initiative. Davis and Hensley (1999) interviewed principals in six California districts and reported that few of them felt that their evaluations were useful and most believed that they reflected political forces beyond their control. Thomas, Holdaway and Ward (2000) collected survey and interview data from principals in Alberta, Canada and found wide variation in school leaders' understandings of the purposes of principal evaluation with many viewing it as ineffective. Reeves (2005) administered a survey to principals across the U.S. and reported that most school leaders felt that principal evaluation did not provide helpful feedback to them and had little impact on their motivation or performance.

Kimball, Milanowski and McKinney (2009) investigated the experiences of two groups of principals in a large district in the western U.S. The district had enacted a standards-based approach to principal evaluation with one group of school leaders while the other group was assessed using the district's traditional approach. Kimball and colleagues found that principals in the standards-based evaluation group were more likely than their counterparts in the second group to indicate that they received useful feedback and that they were satisfied with their evaluation. Goldring and colleagues (2015) explored how 14 principals made sense of and responded to feedback from teachers about their instructional leadership. The authors reported that school leaders experienced cognitive dissonance when their teachers assigned them lower ratings for instructional leadership than the ratings they assigned themselves or when teachers' ratings of their leadership were lower than the average teacher ratings of other principals in their district. This study also found that principals' orientations to feedback were associated with their reactions to it. In particular, these principals often developed a negative and

defensive orientation toward their feedback which hindered their ability to constructively use it for improvement. School leaders with a positive orientation were more likely to be receptive to feedback.

Taken together, these studies indicate that many principals found little utility in their evaluations prior to RTTT, with the exception of school leaders who had a positive orientation to feedback and those who experienced a standards-based approach to evaluation.

A few researchers explored how principal evaluation was associated with principals' leadership practices prior to RTTT. Sun and Youngs (2009) collected survey data from 85 principals in 13 districts in Michigan as well as teachers and district administrators who worked with them. They reported that school leaders were more likely to enact instructional leadership when evaluation focused on such leadership; when the purposes of evaluation included principal development and school reform; and when evaluation addressed leadership in such areas as setting school goals, designing curriculum, monitoring student performance, and promoting teacher development. Hamilton and colleagues (2012) investigated Pittsburgh Public Schools' approach to principal development, which featured summer leadership trainings for school leaders and individual professional growth projects; and evaluation, in which supervisors provided feedback and coaching to principals while also evaluating them. Hamilton et al. found that after the district changed its approach to principal development and evaluation, school leaders spent much more time observing and evaluating teachers and providing them with feedback on instruction; in addition, many principals worked with teachers to analyze student achievement data and took an active role in planning school-based professional development.

In research in Atlanta, New York City, and Oakland, Honig (2012) collected interview data from 162 district administrators, principals, and external support providers. She found that district administrators were more likely to be identified as promoting principals' efforts to strengthen their instructional leadership when they modeled leadership practices, used metacognitive strategies, employed tools to have challenging conversations about instructional leadership, linked tool use to clear definitions of such leadership, differentiated assistance, and connected principals to other sources of support. In an analysis focused on a single district, Honig and Rainey (2014) drew on interview and observation data to investigate the role of district administrators in enhancing professional learning communities for principals (PPLCs). They reported that when district administrators used modeling, tools, differentiation, and brokering, principals were more likely to be generally engaged in PPLC meetings and, in particular, to be engaged in challenging instructional leadership activities such as discussing the implications of leadership practices for instructional quality and student learning.

Goff and colleagues (2014) examined whether feedback and coaching for principals was related to changes in leadership practice as measured by teachers' perceptions. They focused on 52 elementary and middle school principals in a large urban district in the Southeastern U.S. In particular, the authors assigned 26 school leaders to a treatment group which received feedback and coaching and 26 to a control group that only received feedback on their leadership. The authors reported that principals in the treatment group were significantly more likely to make changes in their leadership practice compared to principals in the control group.

In the wake of RTTT, 50 out of 51 states (including the District of Columbia) made changes to principal evaluation policy. Fifty states included measures of leadership and practice in evaluation and forty-six required use of student achievement data. Forty-three states allowed districts to develop their own evaluation systems as long as they were aligned with state policy and thirty-six mandated that all school leaders be evaluated annually (Donaldson et al., 2021a).

Following RTTT, researchers documented ways in which school districts in the U.S. changed their approaches to principal evaluation. In general, districts began to focus primarily on supervising principals and enhancing their professional growth and less on holding them accountable. For example, Anderson and Turnbull (2016) studied six large urban districts that were part of the Wallace Foundation's Principal Pipeline Initiative and found that all six districts included measures of leadership practices and student achievement gains in evaluation and employed evaluation mainly as a strategy to strengthen principals' leadership skills. In addition, all six districts modified the role of principal supervisors to concentrate more on promoting principals' development as instructional leaders and less on monitoring their enactment of district priorities. In a second study, Kimball and colleagues (2015) examined how three large urban districts and two small rural districts restructured their principal evaluation systems. They found that most of the districts had implemented goal setting and continuous improvement cycles for principals. In addition, some of the districts provided formal professional development for principal supervisors while others depended on state training of supervisors.

Goldring and colleagues (2020) studied six large urban districts that were part of the Wallace Foundation's Principal Supervisor Initiative; each district modified the principal supervisor role to focus on enhancing principals' development as instructional leaders. They reported that school leaders began to view their central district offices as more supportive of their schools, their supervisors as focusing on high-quality instruction, and their work with supervisors as helping them become more effective instructional leaders. In particular, many 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” (Goldring et al., 2020, p. 56). In addition, when principals viewed their supervisors as being effective, this was significantly associated with teachers’ perceptions of principal performance.

In another study, Zepeda and colleagues (2014) explored principal evaluation practices in one district post-RTTT. They found that the superintendent confronted a number of challenges including (a) having prior relationships with principals that affected their evaluations of them; (b) rating a principal’s performance in ways that were not consistent with ratings of their school’s performance; and (c) questions related to evaluating principals who were new at schools that had histories of low student achievement.

Taken together, the studies by Goldring et al. (2020), Hamilton et al. (2012), and Sun and Youngs (2009) indicate that principal evaluation systems that focus on instructional leadership, feature feedback from supervisors, and are linked to principal professional development are likely to promote effective instructional leadership practices among principals. But less is known about how principals’ views of the extent to which district principal evaluation includes goal setting, feedback, and professional development (i.e., intrinsic sources of motivation) are associated with whether they perceive that principal evaluation in their districts promotes leadership change and improvement (Donaldson et al., 2021b). Further, there has been little research on how principals experience these aspects of principal evaluation that potentially address intrinsic motivation.

Conceptual Framework

Consistent with our conceptual framework, displayed in Figure 1, we posit that when principal evaluation systems include components that address school leaders’ intrinsic motivation, principals are more likely to perceive such systems as enhancing leadership change and improvement; that is, principals are more likely to view these systems as fair, consistent with their beliefs about effective leadership, and affecting how they apportion time to different leadership activities (H_1). We also posit that when evaluation includes both extrinsic and intrinsic sources of motivation, components that address extrinsic motivation sources (such as the use of evaluation results to hold principals accountable or to make decisions about salaries, promotions, and sanctions) will have a much weaker association with principal perceptions that evaluation promotes leadership change and improvement compared to components that include intrinsic motivation sources (H_2).

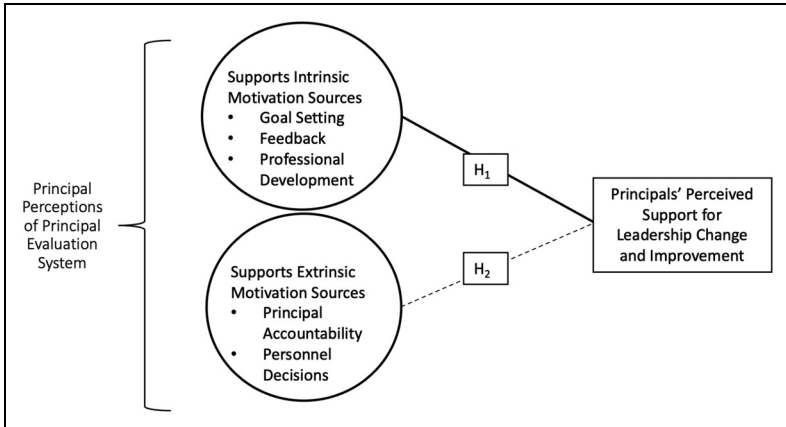


Figure 1. Principal evaluation and sources of motivation conceptual framework.

As noted, we define intrinsic motivation as the act of engaging in behavior because one derives enjoyment or satisfaction from it. Intrinsically motivated behavior is important in two respects. First, it “is performed spontaneously because the person is interested in the behavior itself (Deci et al., 1996, p. 169). For example, a school leader who is intrinsically motivated with regard to instructional leadership will be likely to work with teachers on a regular basis to set learning goals and expectations, monitor their teaching, provide them with feedback, and design professional development to address their instructional needs. Second, intrinsic motivation can lead to lasting changes in leadership practices (Donaldson, 2020). These can include changes in how principals establish and maintain relationships with students, parents, and community members; how they address and make decisions about issues related to equity and diversity; and how they manage finances, school facilities, and personnel.

Goal setting and feedback are closely associated with intrinsic motivation. Goals are most likely to address principals’ intrinsic motivation when they play a direct role in establishing them, when they are precise and challenging, when they are used for self-evaluation, and when there is support from supervisors or others in achieving them (Deci & Flaste, 1995; Donaldson, 2020; Locke & Latham, 2002; Ordóñez et al., 2009). Principals often set goals in such areas as student achievement, school climate, teacher evaluation, allocation of time to various leadership activities, and community relations. Feedback is likely to strengthen intrinsic motivation when it is specific and actionable, when recipients are highly competent, and when it is combined

with guidance for ways to address weaknesses in performance (Deci et al., 1996; Firestone, 2014). Feedback is often provided by principal supervisors, superintendents, and peers.

Principals are likely to experience intrinsic motivation when principal professional development is tied to principal evaluation, when it features goal setting, and when it involves access to supervisors and others who can provide feedback and help them work towards their self-development goals. At the same time, when professional development does not include goal setting or feedback, it is less likely to promote intrinsic motivation. In addition, when districts set clear, coherent expectations over time with regard to principal evaluation, goal setting, feedback, and professional development, school leaders are more likely to experience intrinsic motivation (Firestone, 2014; Honig & Hatch, 2004; Rorrer et al., 2008). Alternatively, when expectations for principal and school performance change frequently and/or are less coherent across evaluation and professional development, they are less likely to experience intrinsic motivation.

As noted, we define extrinsic motivation as the act of engaging in behavior because of an outside expectation, pressure, or incentive to do so. Extrinsic sources of motivation for principals include the use of principal evaluation results to hold principals accountable and make decisions about principal salary increases, promotions, and sanctions. Consistent with our framework, we posit that when evaluation results are used in making personnel decisions, this will have a weak association with principals' perceptions of whether evaluation enhances leadership change and improvement (Donaldson, 2020; Firestone, 2014). That is, we hypothesize that when evaluation is used for such accountability and decision making, principals are less likely to view it as fair, aligned with their vision of effective leadership, or prompting them to make changes in leadership practice (compared to when it features goal setting, feedback, and links to professional development).

In summary, consistent with our conceptual framework, we posit that it is important for principal evaluation to address principals' intrinsic motivation through goal setting, feedback, and professional development because such motivation is often associated with changes in leadership practices and working closely with teachers to promote their development. We also hypothesize that evaluation that includes intrinsic motivation sources is more likely to be perceived as promoting leadership change and improvement than evaluation that addresses extrinsic motivation sources.

Research Methods

We employed a mixed-methods sequential explanatory design, in which we examined quantitative data and then used qualitative data to unpack and

Table 1. Research Questions, Data Sources, and Analytic Strategies.

| Research questions | Data sources | Analytic strategies |
|---|---------------------|----------------------------------|
| 1. What is the association between the degree to which principals report that their districts' principal evaluation systems include intrinsic sources of motivation and extrinsic sources of motivation and their perceptions of whether their districts' evaluation systems promote leadership change and improvement? | Principal survey | Multiple regression analysis |
| 2. How do principals describe and experience intrinsic sources of motivation in the context of their principal evaluation system? | Principal interview | Deductive and inductive analysis |

explain quantitative findings and to extend the range of our inquiry (Creswell & Creswell, 2018; Ivankova et al., 2006; Moss & Haertel, 2016). First, we analyzed quantitative principal survey data. Second, in an intermediate phase, we used findings from the quantitative analysis to inform how we would approach our qualitative sampling and analytic plan. Third, we analyzed qualitative principal interview data. Finally, we connected quantitative and qualitative analyses through the integration of results. The advantage of this design is that it allowed us to gain a more textured and elaborated understanding (Creswell & Creswell, 2018; Ivankova et al., 2006; Moss & Haertel, 2016) of how extrinsic and intrinsic sources of motivation shaped how principals experienced their principal evaluation systems by analyzing multiple data sources in a complementary fashion. Table 1 summarizes our research questions, data sources, and analytic strategies.

Research Phase 1: Quantitative Principal Survey Analysis

In the study's first phase, we used principal survey data to examine the association between principals' perceptions of the extent to which principal evaluation included intrinsic and extrinsic motivation sources and their views of whether it promoted leadership improvement.

Data, Data Collection, Measures. We administered online surveys to principals to understand their perceptions of how their districts evaluated them. We designed the survey items to gather detailed information on principals' experiences with the principal evaluation system in their district, professional development and growth opportunities linked to evaluation, as

well as rewards and sanctions tied to principal evaluation ratings. In addition, we included questions about principals' demographic characteristics (e.g., race/ethnicity, gender), professional experience (e.g., years of experience as principal, years of teaching experience), and school demographic characteristics.

We first constructed principal survey items based on research on principal evaluation, professional development, and extrinsic and intrinsic sources of motivation (Hamilton et al., 2012; Sun & Youngs, 2009). We then revised the survey items based on piloting and in-depth discussions to ensure that the items sufficiently reflected the construct in the measure with relevancy and clarity.

Sample. Our sample included responses from 82 elementary and middle school principals in 21 districts across 3 states (Connecticut, Michigan, and Tennessee). We purposefully chose these focal sites to ensure that our data would reflect considerable variation in state principal evaluation policies, as well as size and demographics of districts. During the years when we collected data for this study (2017–18 and 2018–19), these states granted districts discretion in creating principal evaluation systems that adhered to state guidelines, yet the state policies differed in key ways. For example, these states required that principals' evaluations be based in part on student performance, but weights and what constituted student performance varied. In Connecticut, 45% of the principal's evaluation had to be based on student achievement measures, with 22.5% based on performance on state assessments and 22.5% based on locally-determined indicators of performance aligned to state learning standards. In Michigan, 25% of the evaluation had to be based on student growth and assessment data (this increased to 40% in 2018–19), which included the same data used in teacher evaluations. In Tennessee, 50% of principal evaluation was based on student performance, which included a 35% growth measure and a 15% achievement measure.

We initially administered surveys to 118 principals, and 82 principals responded to the survey, resulting in a response rate of ~70%. To handle missing values in the multiple regression analysis, we used listwise deletion and only included observations with valid values across all variables in the analysis. We selected this approach to utilize complete data for analysis and because missing response cases in our regression models were slightly over 10%, which leaves valid observations that range from 71 to 79 principals out of 82 principals who responded at all across regression models.

Variables. We constructed variables using items from the principal survey. Each individual question was on a 4-point Likert scale that ranged from (1)

Strongly disagree to (4) *Strongly agree*; thus, the average across all of the items for each variable is continuous from 1 to 4. We used multiple items to measure constructs such as extent to which principals reported that their evaluation systems included intrinsic/extrinsic sources of motivation and whether principals perceived principal evaluation as enhancing leadership change and improvement. To gather additional evidence of validity for the measures, we conducted exploratory factor analysis and checked whether items for a certain construct indeed loaded onto the construct. Based on information from the percentage of total variance accounted for by each factor, eigenvalue, and screen plot, we obtained one factor solution for each of the following constructs: extent to which principals reported that principal evaluation addressed intrinsic/extrinsic sources of motivation and perceptions of the degree to which principal evaluation promoted leadership change and improvement.

Across each construct, we selected items that could be well-explained by a factor. Hence, we chose items with factor loading values higher than .3 as such values are viewed as representing a factor clearly. We retained 8 out of 8 items for perceptions of one's principal evaluation system, 4 out of 4 items for extrinsic motivation, and 12 out of 13 items for intrinsic motivation as final items. The Cronbach alpha reliability coefficients for these constructs were 0.90, 0.64, and 0.86, respectively. We used these results for item selection in order to include refined measures in the regression analysis. After we chose items for each construct, we used the average score of each construct in the regression analysis.

The dependent variable was a measure of the degree to which principals believed their district's evaluation system enhanced leadership change and improvement. It was the average score of eight items which asked principals, "To what extent do you agree with the following statements about principal evaluation in your district?" These items included: "Principal evaluation is fair to all principals in my district, regardless of the type of school in which they work," "My district's principal evaluation system aligns with my vision of what makes a good principal," and "I have altered how I allocate my time toward the leadership tasks emphasized in the principal evaluation system." We used principals' responses to these items to create a measure of the degree to which they perceived that principal evaluation promoted leadership change and improvement. The Cronbach alpha reliability coefficient for these items was $\alpha = 0.90$. See Table 2 for the full list of the items that made up this dependent variable and the two independent variables and Table 3 for descriptive statistics for all three variables.

We included two independent variables of interest, one consisting of a mean score variable for four extrinsic motivation items and one consisting of a mean score variable for twelve intrinsic motivation items. The stem for the extrinsic motivation items was as follows: "In your district, how important

Table 2. Variable Description.

| Variable | Levels of measurement | Survey items |
|--|---|---|
| Principal's perception of degree to which principal evaluation promotes leadership improvement | 4-point Likert-type scale (1) strongly disagree (2) disagree (3) agree (4) strongly agree | To what extent do you agree with the following statements about principal evaluation in your district? 1) Principal evaluation is fair to all principals in my district, regardless of the type of school in which they work. 2) Principal evaluation is fair to all principals in my district, regardless of their personal characteristics (e.g., race, gender, sexual orientation). 3) The principal evaluation system does a good job of distinguishing between effective and ineffective principals. 4) I have altered how I allocate my time toward the leadership tasks emphasized in the principal evaluation system. 5) I have a clear understanding of what I need to do in order to achieve high rating in the evaluation system. 6) Principal evaluation has helped me think about my leadership strengths and weaknesses. 7) Principal evaluation is fair to all principals in my district regardless of their evaluator. 8) My district's principal evaluation system aligns with my vision of what makes a good principal. |
| Sources of motivation | 4-point Likert-type scale (1) not at all (2) some extent | Please indicate how principal evaluation and principal professional development are related in your district. 1) Professional development sessions are offered to address issues identified in principal evaluations. 2) Individual action plans are developed to overcome issues identified in principal |

(continued)

Table 2. (continued)

| Variable | Levels of measurement | Survey items |
|---|--|--|
| (3) moderate extent (4) great extent | <p>evaluations.</p> <p>3) Supervisors help principals accomplish their self-development goals through the evaluation process.</p> <p>4) Experts from within the district are assigned to help principals improve in certain areas.</p> <p>5) Experts from outside the district are assigned to help principals improve in certain areas.</p> <p>6) Principals' self-evaluation is part of district evaluation program.</p> <p>Please indicate the extent to which each of the following applies in your district:</p> <p>7) The intended purposes of evaluation do not include professional development.^a</p> <p>8) The evaluation procedures and structures do not collect information for professional development.^a</p> <p>9) Lack of time prevents use of evaluation results for professional development.^a</p> <p>10) Lack of money prevents use of evaluation results for professional development.^a</p> <p>11) I do not have formal professional learning and growth/professional development in my district.^a</p> <p>12) My evaluation results don't match the areas where I feel I need to improve as a leader.^a.</p> | |
| Extrinsic motivation sources | <p>4-point likert-type scale</p> <p>(1) not at all important</p> <p>(2) a little important</p> <p>(3) somewhat important</p> <p>(4) very important</p> | <p>In your district, how important are the following purposes of principal evaluation?</p> <p>1) To provide information for use in making decisions about hiring, principal promotion, (re)assignment, and contract renewal within the district</p> <p>2) To provide evidence to determine merit salary increases</p> <p>3) To provide evidence to determine sanctions for principals</p> <p>4) To hold principals accountable for student achievement</p> |

(continued)

Table 2. (continued)

| Variable | Levels of measurement | Survey items |
|--|---|---|
| Leadership areas addressed in the principal evaluation system | | |
| Managerial leadership | 4-point likert-type scale (1) not at all (2) some extent (3) moderate extent (4) great extent | To what extent are each of the following areas of leadership addressed in your district's principal evaluation system?: -Managerial leadership (e.g., human resources, finances, school facilities) |
| Personal attributes | 4-point likert-type scale (1) not at all (2) some extent (3) moderate extent (4) great extent | To what extent are each of the following areas of leadership addressed in your district's principal evaluation system?: -Personal attributes (e.g., communication skills, ethical behavior, decision making) |
| Instructional leadership | 4-point likert-type scale (1) not at all (2) some extent (3) moderate extent (4) great extent | To what extent are each of the following areas of leadership addressed in your district's principal evaluation system?: -Instructional leadership (e.g., implementation of vision, data-based decision making, teacher professional development, supporting and/or monitoring instruction) |

(continued)

Table 2. (continued)

| Variable | Levels of measurement | Survey items |
|--|---|--|
| Principal background Total years of principal experience | Ordinal variable (1) less than 1 (2) 1–3 (3) 4–8 (4) 9–12 (5) 13 or more | How many years have you worked as a principal (including 2017–18)? |
| Male | Binary variable (0) female (1) male | Are you male or female? |

^aReverse coded items.

Table 3. Descriptive Statistics for Variables Used in Multiple Regression Analysis.

| Variable | Mean | Standard deviation |
|--|------|--------------------|
| Degree to which evaluation system promotes leadership improvement | 2.90 | 0.61 |
| <u>Sources of motivation</u> | | |
| Extrinsic motivation | 2.41 | 0.70 |
| Intrinsic motivation | 2.83 | 0.57 |
| <u>Areas addressed in the principal evaluation system</u> | | |
| Managerial leadership | 2.88 | 0.86 |
| Personal attributes | 3.15 | 0.86 |
| Instructional leadership | 3.72 | 0.59 |
| <u>Principal and school background characteristics</u> | | |
| Total years of principal experience | 3.34 | 1.05 |
| Male | 0.43 | 0.50 |
| Percentage free/reduced price lunch students | 0.35 | 0.22 |

Note. Total years of principal experience is an ordinal variable. (1): less than 1 year, (2) 1–3 years, (3) 4–8 years, (4) 9–12 years, (5) 13 or more years.

are the following purposes of principal evaluation?" These four items included: "To provide information for use in making (personnel) decisions," "To provide evidence to determine merit salary increases," "To provide evidence to determine sanctions for principals," and "To hold principals accountable for student achievement." The Cronbach alpha reliability coefficient for these items was $\alpha = 0.64$.

The stem for the 12 intrinsic motivation items was as follows: "Please indicate how principal evaluation and principal professional development are related in your district." These items included: "Principals' self-evaluation is part of district evaluation program," "Supervisors help principals accomplish their self-development goals through the evaluation process," "Individual action plans are developed to overcome issues identified in principal evaluations," "The intended purposes of evaluation do not include professional development (reverse coded)," and "My evaluation results don't match the areas where I feel I need to improve as a leader" (reverse coded). The Cronbach alpha reliability coefficient for these items was $\alpha = 0.86$.

We included three control variables that asked about the extent to which principals perceived that three areas of leadership (i.e., managerial leadership, personal attributes, instructional leadership) were addressed in their districts' evaluation systems. We also included controls for principals' backgrounds, such as gender and total years of principal experience, and school characteristics including the percentage of students eligible for free or reduced-price lunch. These variables are described in Table 2 and descriptive statistics are presented in Table 3.

Analytic Approach. We fit multiple regression models to identify factors that were associated with principals' perceptions of their districts' evaluation systems. In Models 1 and 2, we examined the association between the degree to which principals reported that evaluation included intrinsic and extrinsic sources of motivation, respectively, and principals' perceptions of whether evaluation enhanced leadership change and improvement. In Model 3, we included both sources of motivation variables. In Model 4, we incorporated one set of controls, the perceived focus areas of evaluation. Finally, in Model 5, we incorporated all of the variables: the two sources of motivation variables, controls for leadership focus of evaluation, and controls for principals' backgrounds and school characteristics. Model 5 is our final model and specification for this model is represented in equation (1).

$$y_i = \beta_0 + \beta_1(\text{Extrinsic}_i) + \beta_2(\text{Intrinsic}_i) + \vec{\beta}_3(\vec{\text{LeadArea}}_i) + \vec{\beta}_4(\vec{\text{PrinSchControls}}_i) + \varepsilon_i \quad (1)$$

Research Phase 2: Qualitative Analysis of Principal Interview Data

In the second phase of our study, we used qualitative principal interview data to more fully explore principals' experiences with intrinsic sources of motivation embedded in their districts' principal evaluation systems and explain variation in how school leaders experienced such sources with regard to motivating leadership change and improvement.

Sample. For our larger study, we conducted interviews with each principal who completed our survey. From this larger sample, we purposefully selected six principals from among the 82 survey participants to create a maximum variation sample for the qualitative phase of our study (Creswell & Guetterman, 2019). We chose three of these school leaders from among the 11 survey participants who in the principal survey (a) rated the degree to which their districts' principal evaluation systems addressed intrinsic sources of motivation most highly and (b) perceived that these evaluation systems promoted leadership change and improvement (referred to as the first group). We selected the other three school leaders from among the 10 survey participants who in the principal survey (a) rated the degree to which principal evaluation addressed intrinsic sources of motivation at the lowest levels and (b) did not view evaluation as promoting leadership change or improvement (referred to as the second group).

By selecting a subsample of participants in this fashion, we were able to capture rich, layered data regarding principals' perceptions of how their district's evaluation system incorporated intrinsic motivation sources and whether/how they saw this connected to the potential for the system to enhance leadership change and improvement. See Table 4 for information on the qualitative sample. In the interview sample, we included four principals from Michigan and two from Connecticut; and four elementary school principals and two middle school principals. Their levels of school leadership experience ranged from 3 to 20 years. Their districts ranged in size and location from small rural districts to mid-sized suburban districts.

Data. During 2017–18 and 2018–19, we conducted semi-structured interviews with principals in order to better understand their perceptions of principal evaluation in their districts. We asked them about how their districts helped them develop their leadership skills, principal evaluation structures and processes, the types of leadership activities emphasized, their perceptions of strengths and weaknesses of the evaluation systems, and whether and how principal evaluation prompted them to shift their leadership priorities and practices. Principal interviews lasted 60 to 75 min. We recorded the

Table 4. Principal, School, and District Characteristics for Principal Interview Sample.

| Principal pseudonym | Years of principal experience (current school, overall) | School level | School characteristics | District characteristics | State |
|--------------------------------|--|--------------|--|--|-------|
| First group principals | | | | | |
| Mr. Means | 6 years, 8 years | Elementary | 460 Students 50% FRL 15% Students of color | 2,140 Students 45% FRL 15% Students of color | MI |
| Ms. Mussina | 10 years, 20 years | Elementary | 305 Students 10% FRL 15% Students of color | 3,375 Students 10% FRL 15% Students of color | CT |
| Mr. Palmer | 11 years, 11 years | Middle | 900 Students 35% FRL 35% Students of color | 5,455 Students 40% FRL 35% Students of color | MI |
| Second group principals | | | | | |
| Mr. Ripken | 9 years, 9 years | Elementary | 390 Students 40% FRL 20% Students of color | 5,110 Students 25% FRL 15% Students of color | MI |
| Ms. Robinson | 3 years, 3 years | Middle | 665 Students 55% FRL 15% Students of color | 2,835 Students 55% FRL 10% Students of color | MI |
| Ms. Weaver | 5 years, 5 years | Elementary | 205 Students 25% FRL 35% Students of color | 1,175 Students 25% FRL 30% Students of color | CT |

Note. We have rounded student enrollment, percent FRL, and percent students of color so as to protect anonymity of our participating schools and districts. FRL = Qualifies for free or reduced-price lunch.

interviews and transcribed them verbatim. We used pseudonyms for all principal interview participants.

Analytic Approach. For each principal interview, we wrote a detailed analytic memo immediately following the interview that described the tone and meaning discerned at the time of the interview. We created initial codes based on theory and research on extrinsic and intrinsic motivation. We generated subsequent codes while analyzing data from the interviews (see Table 5 for our lists of initial and final codes). By grouping together categories and using the constant comparative method (Strauss & Corbin, 1990), we moved to higher levels of abstraction and eventually derived the following codes: autonomy, district expectations, extrinsic sources of motivation, feedback, interactions with district administrators, intrinsic sources of motivation, principal professional development, and principal evaluation (Miles et al., 2020; Strauss & Corbin, 1990). Once the final codes were established, two research team members separately coded all six interviews. We compared these separate efforts and in cases of disagreement, we discussed the codings until consensus was reached (Hill et al., 2005).

For the second stage of qualitative data analysis, we compiled case reports by principals’ survey responses (i.e., the degree to which they reported that evaluation included intrinsic motivation sources and whether they felt that it promoted leadership change and improvement) and identified emergent themes regarding (a) principal interactions with district administrators and principal professional development; and (b) autonomy, feedback, and district expectations. Through this process, we identified several linkages among these key codes and themes.

Table 5. Initial and Final Codes.

| Initial codes | Final codes |
|---|---|
| Autonomy | Autonomy |
| District expectations | District expectations |
| Expectancy | Extrinsic motivation |
| Extrinsic motivation | Feedback |
| Feedback | Interactions with district administrators |
| Interactions with district administrators | Intrinsic sources of motivation |
| Intrinsic sources of motivation | Principal professional development |
| Principal professional development | Principal evaluation |
| Principal evaluation | |
| Self-efficacy | |
| View self as learner | |

As we ascertained and described the processes involved in evaluating and enhancing principal leadership in this study, we determined that the principal was the key analytical unit. Thus, our data analysis involved looking across the learning opportunities experienced by and expectations placed on the two groups of principals. When it became evident that some principals reported having constructive interactions with district administrators, receiving useful feedback from them, and experiencing consistent district expectations over time, we created additional data displays to confirm these patterns, while remaining attentive to disconfirming evidence (Miles et al., 2020). In this way, we were able to analyze the ways in which extrinsic motivation addressed sources of intrinsic motivation through principal evaluation for some school leaders, but not for others.

Procedures for Establishing Validity of Interview Data. In this study, we took multiple steps to establish the validity of the interview data reported on in this manuscript. These included use of multiple data collection methods, a multiple case design, and peer review and debriefing (Miles et al., 2020; Strauss & Corbin, 1990; Yin, 2018). We employed multiple data collection methods to strengthen the validity of the interview data (Yin, 2018). In particular, we drew on the surveys of the six interview participants to check on their experiences with principal evaluation and professional development. We found that the participants' responses to interview questions about these topics were nearly identical to their survey responses concerning the same topics.

Second, by including two distinct groups of principals in the interview sample, we incorporated a multiple case design featuring replication logic (Yin, 2018). More specifically, by including school leaders who did/did not feel that principal evaluation included intrinsic motivation sources and/or that it promoted leadership change and improvement, we were able to test our theory that the two groups of principals experienced different sources of intrinsic motivation. Finally, we received external feedback on our research design and initial findings from researchers in the areas of principal leadership and evaluation (Strauss & Corbin, 1990). In particular, these colleagues encouraged us to (a) restrict the focus of the qualitative analysis to two groups of three principals each; and (b) draw on prior research on teacher motivation and evaluation to inform our conceptual framework.

Findings

We begin by describing quantitative results regarding the association between principals' perceptions of the extent to which principal evaluation systems

included intrinsic sources of motivation and extrinsic sources of motivation and their perceptions of whether evaluation enhanced leadership change and improvement. We then turn to qualitative findings regarding principals' experiences with intrinsic sources of motivation in the context of principal evaluation.

Association Between Perceptions of Extrinsic and Intrinsic Motivation Sources and Whether Evaluation Systems Promote Leadership Change and Improvement

In Table 6, as a baseline, Models 1 and 2 reported the association between the extent to which principals reported that their district principal evaluation systems included extrinsic and intrinsic sources of motivation, respectively,

Table 6. Factors Predicting Principals' Perceptions of Evaluation Systems.

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
|---|----------------|----------------|----------------|----------------|----------------|
| Sources of motivation | | | | | |
| Extrinsic motivation | .36* (0.10) | | .13 (0.08) | .04 (0.08) | .02 (0.09) |
| Intrinsic motivation | | .64* (0.09) | .61* (0.10) | .52* (0.11) | .48* (0.11) |
| Leadership areas | | | | | |
| Managerial leadership | | | | .04 (0.07) | .06 (0.07) |
| Personal attributes | | | | .22* (0.08) | .18 (0.08) |
| Instructional leadership | | | | .09 (0.10) | .17 (0.11) |
| Principal and school backgrounds | | | | | |
| Total years of principal experience | | | | | -.02 (0.06) |
| Female | | | | | .15 (0.11) |
| % of free/reduced price lunch students | | | | | -.12 (0.27) |
| N | 76 | 79 | 76 | 75 | 71 |
| R ² | 0.13 | 0.41 | 0.44 | 0.51 | 0.53 |

Note. Coefficients are standardized. Standard errors in parentheses.

*p < .05.

and their perceptions of whether their evaluation systems promoted leadership change and improvement. When each source of motivation was incorporated independently, we found a positive and statistically significant association between (a) the degree to which principals reported that evaluation addressed extrinsic sources of motivation ($\beta = .36, p < .05$, Model 1) and (b) the degree to which they reported that it addressed intrinsic sources ($\beta = .64, p < .05$, Model 2) and the degree to which they reported that it enhanced leadership improvement. Consistent with our conceptual framework, we found that principals' perceptions of the degree to which evaluation promoted leadership change and improvement were strongly associated with the degree to which they reported that it included intrinsic motivation sources.

While principals' perceptions of the degree to which evaluation included extrinsic sources of motivation were also associated with principals' perceptions of whether their evaluation system enhanced leadership change and improvement, we found that the magnitude of the coefficient was approximately half that of the intrinsic motivation measure. Model 3 incorporated principals' perceptions of the degree to which evaluation addressed both extrinsic and intrinsic sources of motivation simultaneously. When both measures were included, we found that principals' perceptions of the degree to which evaluation addressed intrinsic motivation remained positively and significantly associated with principals' perceptions of the degree to which evaluation promoted leadership change and improvement ($\beta = .61, p < .05$, Model 3), whereas the coefficient for principals' perceptions of the degree to which evaluation addressed extrinsic motivation was no longer statistically significant and the magnitude attenuated substantially from Model 1.

In Model 4, we built upon Model 3 by adding controls for the extent to which the principal evaluation system focused on three key leadership areas (i.e., managerial leadership, personal attributes, instructional leadership) and exploring whether the presence of these elements might otherwise explain the perceived relationship (evident in the first three models) between the degree to which principals reported that evaluation included intrinsic sources of motivation and their perceptions of the degree to which their evaluation system enhanced leadership change and improvement. We found that principals' perceptions of the degree to which evaluation addressed intrinsic sources of motivation continued to have the strongest positive association with perceptions of whether evaluation promoted leadership change and improvement ($\beta = .52, p < .05$, Model 4).

We found that one control variable, personal attributes (e.g., communication skills, ethical behavior, decision making), addressed in the principal evaluation system was also positively associated with principals' views about whether evaluation promoted leadership change and improvement ($\beta = .22$,

$p < .05$, Model 4). Our final model indicated that the positive association between the degree to which principals reported that evaluation included intrinsic sources of motivation and principals' perceptions of the degree to which their evaluation system enhanced leadership change and improvement ($\beta = .48$, $p < .05$, Model 5) persisted after controlling for principals' backgrounds and school characteristics. But the association between focusing on personal attributes in principal evaluation was no longer significant and importantly was attenuated almost entirely to zero, indicating that the significant relationship between personal attributes and the principal perception in Model 4 is fully explained by these controls. Taken together, our findings suggest that the changes in coefficients for extrinsic motivation across models are not just due to the precision of our estimates, but the inclusion of other controls and predictors meaningfully accounted for the variation in perceptions that we might have otherwise attributed to extrinsic motivation sources.

From these analyses, we conclude that the degree to which principals reported that principal evaluation included intrinsic sources of motivation appeared to be consistently associated with principals' perceptions of whether evaluation promoted leadership change and improvement. Across all models, we found that the association between this predictor variable and the outcome variable remained strong and statistically significant. This suggests that when principals perceive evaluation as emphasizing intrinsic sources of motivation, such as linking evaluation to professional development, it is more likely to lead to leadership change and improvement.

Understanding How Principals Experience Intrinsic Motivation Sources in the Context of Principal Evaluation

We draw on interview data with the six focal interview participants to describe principals' experiences with different aspects of principal evaluation and to examine how school leaders experienced aspects of evaluation such as feedback, autonomy, and district expectations that potentially addressed intrinsically-motivated behavior. As a reminder, in their responses to our principal survey three of these principals rated the degree to which their districts' principal evaluation systems included intrinsic sources of motivation very highly and felt that these systems enhanced leadership change and improvement while the other three school leaders did not feel that that evaluation addressed intrinsic sources of motivation and did not view it as promoting leadership change and improvement.

While both groups reported in interviews similar expectations with regard to setting personal goals and attending principal meetings in their districts, there were clear differences between groups with regard to their interactions

with district administrators (i.e., supervisors), the nature of principal professional development activities in which they participated, and their experiences with feedback, autonomy and district expectations. These differences help to explain the results of our regression analyses, including variation between the two groups in their perceptions of the degree to which principal evaluation promoted leadership improvement.

Goal Setting and Principal Meetings

To some extent, the six principals in the focal interview sample reported having similar experiences with principal evaluation. In particular, all principals in both groups reported setting personal goals as part of principal evaluation, working on growth plans, and/or using data to document their performance as school leaders. For example, Mr. Palmer, a middle school principal from Michigan in the first group, explained that his district used the Marzano School Leader Evaluation Model. For principals, this model involved completing a growth plan at the start of each school year and then documenting ways that they were addressing several of the 23 elements in the model during that school year. Palmer noted that the Marzano model was “systematized in common language and systematized where we’re all reading the same information about how you should be scored.” Similarly, Ms. Weaver, an elementary school principal from Connecticut in the second group described the role of goal setting in her district’s evaluation system: “I would say the strengths are it gives you a place to start at the beginning of the year and goals to aspire to. It’s a formal record of strategies that you’re going to use to improve, so all of that’s due at the beginning of the year.”

Both groups of principals also indicated in interviews that they attended principal meetings coordinated by their districts on a regular basis. Ms. Mussina, an elementary school principal from Connecticut in the first group, reported that she met with other school leaders in her district weekly to discuss their leadership experiences as well as their experiences with teacher and principal professional development and evaluation. For his part, Mr. Means, an elementary principal in the first group, explained that he attended biweekly principal meetings with his counterparts in his district and monthly meetings with school leaders from across his county. Similarly, Mr. Ripken, an elementary principal from Michigan in the second group, stated, “Every Tuesday, all the administrative staff in the district are out of their buildings for two hours to have a meeting.” During these meetings, principals discussed such topics as student testing, teacher professional development, and teacher evaluation.

Interactions with District Administrators and Principal Professional Development

Despite both groups of principals reporting similar experiences with regard to goal setting and principal meetings, there were clear differences between the groups with regard to their interactions with district administrators (i.e., supervisors) and the nature of the principal professional development activities in which they participated. All three principals in the first group reported having regular, productive interactions with their supervisors and that principal evaluation was closely tied to their own professional learning and growth. For example, Mr. Palmer met twice a month with a district administrator to review his progress towards the goals he identified for himself at the start of the school year. In his words, “She and I are spending time talking about those every two weeks—my time with her is always about things that are already documented that I’m doing or working toward in this evaluation over here so they don’t seem separate.” Mr. Means explained that due to his district’s small size (it had only four schools), his superintendent visited his school at least once a week to check in and talk with him about his various responsibilities. For her part, Ms. Mussina explained that she met frequently with her superintendent and that these meetings were closely tied to the evaluation process: “She added, “He just did a cycle of observing me actually ... he wanted to be part of (an observation) cycle I was doing with a teacher. So he observed me with the pre, the during, and the post.”

In contrast to the principals in the first group, none of the focal interview participants in the second group reported having frequent, productive interactions with district administrators. According to Ms. Weaver, “I used to look forward to the visits (from my superintendent) and now, I can’t say that I look forward to the visits. I sort of don’t really, I guess I wonder what mood she might be in. And how much that’s going to affect what happens during the observation.” For her part, Ms. Robinson, a middle school principal from Michigan in the second group, explained that she would have liked to meet with her superintendent to discuss her goals as a principal, her performance, and her strengths and weaknesses, but she had few opportunities to do so. Mr. Ripken noted that his view of his role as principal and how leaders should be evaluated differed from that of his superintendent who focused much more on consequences. He noted that when he made mistakes, his superintendent would set up formal meetings with him and document his mistakes in writing. In his words, “There’s a formality to it that had never been the culture prior to his arrival. It had always been, you know ‘We’ll talk about it. Let’s fix it.’ It was always about learning and growth, not so much about documenting and consequence.”

There were also differences between the two groups of principals with regard to opportunities for professional development linked to principal evaluation. Mr. Means, Ms. Mussina, and Mr. Palmer all met regularly with other principals in their districts to discuss work-related issues. In Mr. Palmer's words, "The amount of meetings, the amount of time that we have to talk about our systems and to help us grow we feel it growing, we're growing together and so that's been a huge benefit." Unlike these three principals, the school leaders in the second group typically engaged in principal professional development on their own and had few opportunities to collaborate with their counterparts in their districts. For example, Mr. Ripken attended the state elementary principals' conference each year at his own volition. While he valued opportunities to make connections with principals from outside his district, he explained that principals within his district had little contact with each other and rarely worked together on common initiatives or supported one another as they worked to improve their leadership.

Feedback, Autonomy, and District Expectations

In addition to identifying differences between the two groups of principals involving interactions with district administrators and principal professional development activities, we found clear differences between them in the nature of the feedback that they received, how they made use of autonomy that they experienced, and the degree to which their districts maintained consistent expectations over time. We found that feedback, autonomy, and district expectations all seemed to affect internal motivation; variation in principals' experiences with these factors seemed associated with differences in their experiences with principal evaluation and its impact on their practices.

We found notable differences between the two groups of principals in the feedback they received; principals in the first group reported receiving routine and extensive feedback through their principal evaluation system and they reported valuing it greatly. For example, Ms. Mussina met with her superintendent twice a month and received regular feedback on issues related to curriculum, instruction, teacher evaluation, budgeting, and situations with teachers, students, and families. According to her, "I'm a learner and they help me grow my learning too." Mr. Means also valued the feedback that he received from district administrators: "I go looking for feedback because I think it's probably one of the most important things, how do you know if you're growing or not if you don't get that exchange, you know of real honest input?"

Mr. Palmer also appreciated opportunities to meet with district administrators about his leadership performance. He stated, "I value that when someone

makes time for me to sit down and listen to the things I'm working on. Here's what I'm struggling with. Here's the growth that we had—the fact that they (provide in-person feedback) that's a positive consequence. That really means something.” The three principals in the first group valued feedback that promoted their own individual growth as school leaders and the growth of their schools. In addition, they appreciated having regular meetings with supervisors and other district administrators during which they were able to discuss their work and receive feedback on it. Further, they seemed to value feedback because it enhanced their sense of their own competence.

On the other hand, principals in the second group reported receiving little feedback on their performance and/or did not value the feedback that they received. Ms. Robinson explained that her superintendent provided brief, written feedback on her annual evaluation ratings, but did not meet with her, provide feedback on leadership practices, or suggest ways for her to improve. In Ms. Weaver's case, she received a detailed formal written evaluation from her superintendent but did not have an opportunity to meet with her about it. In addition, when her superintendent did provide oral feedback, Ms. Weaver felt it was not connected to the evaluation process. For his part, Mr. Ripken disagreed with his superintendent's exclusive focus on increasing students' scores on state standardized tests at the expense of other leadership responsibilities. In his words, “I decided, not that I don't value feedback, I still do. I just don't value this particular group of administrators' feedback ... I didn't feel like we were all on the same team.”

In the interviews, principals in both groups described having a good deal of autonomy with regard to carrying out their roles. But school leaders in the first group were encouraged by supervisors to use their autonomy to address issues related to instruction and student learning while those in the second group were rarely provided with similar direction and used it to focus on students' personal needs. For example, Mr. Means described taking risks to explore new ways of supporting teachers and students at his school. According to him, he felt comfortable doing so “because I feel like there's professional trust” (with district administrators). “I can take risks if they're calculated, not just on a whim.” Similarly, Ms. Mussina felt empowered by her superintendent to have difficult conversations with her teachers about their performance and the need for instructional improvement. For his part, Mr. Palmer reported having freedom to make decisions about instruction and student learning in the ways that he saw fit: “You have a ton of influence about decision making. That's really critical and you have a lot of influence on how you get people to move a whole organization in the direction you want it to go. That's powerful.”

While the principals in the second group also experienced a certain amount of autonomy, they reported employing it to focus more on student needs than

instruction or learning. For example, when asked what he spent most of his time on as principal, Mr. Ripken explained that he primarily concentrated on helping students and families in need. In his words, "I don't think anything takes precedence over supporting kids in crisis." Ms. Robinson also reported focusing on students' needs. She noted that a top priority for her was "being able to deal quickly and effectively with student issues, to find resources and assistance for kids who need them, to know which community agencies to reach out to, communication with families regarding (these) issues." For her part, Ms. Weaver described engaging in a number of non-instructional activities with students including sitting with students during lunch, interacting with them during recess, and addressing misbehavior. Second group principals reported much less direction provided by their evaluation system in terms of how they focused their time, which often seemed to result in a focus on individual student needs instead of organizational and instructional improvement.

We found that the two groups of principals' experiences with feedback and autonomy seemed related to district principal evaluation expectations regarding leadership in their districts. The school leaders in the first group reported experiencing clear, coherent expectations for leadership practice over time. For example, Mr. Means stated that his superintendent had consistently placed a strong emphasis on instructional leadership: "One of the key things I knew, when he came here seven, eight years ago is trying to shift principal role from manager to instructional leader ... So, I try to put number one as instruction. How do we make sure we can hit our building goals and even district goal of creating readers and problem solvers." Mr. Means added that the superintendent's approach to principal and teacher evaluation had been consistent over time as well. For her part, Ms. Mussina explained that her superintendent had always valued the contributions of principals and that the district's principal evaluation system "reflects the real work, my real work and the work of the school." Mr. Palmer reported that in his district, principals had always faced rigorous expectations: "You manage a building, you lead PD, you're always facilitating this, Professional Growth, the School Improvement process it's a little bit, it's more involved than in some other districts."

In contrast, the school leaders in the second group did not report experiencing clear, coherent evaluation expectations regarding leadership practice over time. For example, Ms. Robinson explained that principals' responsibilities in the area of teacher evaluation in her district had recently increased: "Our leadership has changed a little bit in the last couple of years and so now, we're given far more tasks to do than we ever were before." Ms. Weaver explained that her monthly meetings with her superintendent had become more

evaluative than previously; she described her superintendent as “very focused and she’s very driven.” In the case of Mr. Ripken, he had received strong evaluations for several years, but then his superintendent shifted expectations to place much more emphasis on student proficiency on state tests. The superintendent asked him to meet “right after our state proficiency scores had come out ... I was just really blown away by the conversation because it was completely out of the blue ... I personally had always valued growth over proficiency and I can argue that all day long, but I wasn’t really given the opportunity to argue that and that was just the new lens.”

Discussion

In this mixed methods study, we drew on survey data from 82 elementary and middle school principals from three states to examine factors associated with the degree to which they perceived their district principal evaluation systems to promote leadership change and improvement. We found that when school leaders reported that their evaluation systems included intrinsic sources of motivation, they were likely to perceive their evaluation systems as enhancing leadership change and improvement. In contrast, when principals indicated that their evaluation systems addressed extrinsic sources of motivation (without also addressing intrinsic sources), we found that they were unlikely to view these systems as bolstering leadership improvement. In this study, we also used interview data from 6 of the 82 principals to investigate differences among school leaders in their experiences with principal evaluation. In this section, we discuss our main findings in relation to other research on educator evaluation and development, note some limitations of our analysis, and discuss implications for future research.

Theory and prior research suggest that personnel evaluation systems that place a strong emphasis on extrinsic sources of motivation such as high-stakes accountability and financial incentives can undermine intrinsic sources of motivation for educators and employees in other professions (Deci et al., 1996; Donaldson, 2020; Firestone, 2014). In this study, we carried out quantitative analyses that built on this prior work. We found that when principals perceived their district principal evaluation systems to emphasize extrinsic motivation without also including intrinsic sources of motivation, they were unlikely to view these systems as promoting their own efforts to strengthen their leadership practices. On the other hand, when school leaders reported their district principal evaluation systems as addressing intrinsic sources of motivation, we found that they were more likely to view these systems as reinforcing leadership change and improvement. These findings are consistent with research by Honig who reported that principals were more likely to

view district administrators as encouraging their attempts to augment their instructional leadership when they employed modeling, metacognitive strategies, tools, differentiation, and brokering; and connected school leaders to other sources of support (Honig, 2012; Honig & Rainey, 2014).

Sources of intrinsic motivation in this study included reporting that self-evaluation was part of district principal evaluation, principal professional development and supervision were available to address issues identified in evaluation, and evaluation purposes and procedures were connected to professional development. Our findings are consistent with other studies that have shown that principal evaluation systems that include feedback from supervisors and are closely tied to professional development are likely to lead to effective leadership practices among principals (Goldring et al., 2020; Hamilton et al., 2012; Sun & Youngs, 2009).

In our qualitative analyses, we found that for the three principals in the first group, principal evaluation provided extrinsic sources of motivation that addressed several sources of intrinsic motivation: interactions with district administrators, principal professional development, goal setting, feedback, and district expectations (Deci et al., 1996). Mr. Means, Ms. Mussina, and Mr. Palmer reported meeting frequently with district administrators to discuss their own leadership activities, and typically using the principal evaluation process to frame these interactions. In addition, all three principals participated in principal meetings and other professional development activities with other school leaders in their districts. Mr. Means, Ms. Mussina, and Mr. Palmer all experienced a certain level of autonomy in carrying out their leadership roles and all received regular, constructive, in-person feedback on their practices from district administrators through the principal evaluation process (Firestone, 2014). Further, district expectations regarding leadership remained consistent over time in these principals' districts.

For their part, the three principals in the second group—Mr. Ripken, Ms. Robinson, and Ms. Weaver—experienced some of the same potential sources of intrinsic motivation as those in the first group, but these experiences were not addressed by their districts' evaluation systems. These three school leaders interacted with district administrators and participated in professional development, but these activities were based primarily on their individual interests (as opposed to collective participation or being based on principal evaluation) and none of them reported having productive interactions or receiving useful feedback from district administrators in the context of evaluation. Mr. Ripken, Ms. Robinson, and Ms. Weaver had a degree of autonomy, but did not use it to support instruction or student learning. Finally, expectations regarding principal leadership and evaluation in their districts changed over time, which potentially further weakened their intrinsic motivation (Firestone, 2014).

There were some limitations in this study. First, we included elementary and middle school principals from small and mid-sized districts in Connecticut, Michigan, and Tennessee; thus, our findings generalize to similar districts in those states, but not necessarily to larger districts or to districts in other states. Thus, one recommendation for future research would be to examine district principal evaluation systems and school leaders' experiences in larger districts and additional states. Second, our survey sample was somewhat small; the initial survey sample was 82 and the sample in the final model was 71 principals. The size of the survey sample limits our ability to generalize our findings to a larger population. Therefore, we recommend that future research include larger survey samples (e.g., $n = 100$ principals or more). Third, in our qualitative sample, both the first group and the second group included (a) principals at the elementary and middle school levels and (b) principals who worked in low- and high-poverty schools. But we were not able to investigate how other contextual factors (e.g., school or district size, principal experience) might have affected school leaders' experiences with principal evaluation systems. Thus, a third direction for future research would be to explore how such factors may influence principals' experiences with evaluation.

Fourth, in this study, we identified a significant association between principal evaluation systems that included intrinsic sources of motivation and principal perceptions that evaluation promoted leadership change and improvement, but we were not able to establish a causal relationship between these variables. Therefore, a fourth suggestion for future scholarship would be to test interventions that focus on intrinsic motivation sources (e.g., goal setting, feedback, and professional development) to see if they are causally related to principals' perceptions of evaluation and outcomes such as leadership practice changes and improved student achievement. Finally, in the qualitative portion of this study, we relied on interviews with principals to examine their experiences with principal evaluation. In future work, we recommend that researchers also collect data from district administrators, teachers, and others to more fully understand implementation of principal evaluation and professional development, and how they potentially shape leadership practices and teachers' experiences.

In summary, there have been significant changes in principal evaluation policies and practices during the past 10 years, but there is a need to consider principals' experiences with evaluation. In this study, we explored school leaders' perceptions of district evaluation systems; our findings provide evidence that evaluation systems that feature sources of intrinsic motivation are more likely than other systems to help principals engage in leadership change and improvement.

Acknowledgements

We are grateful to the educational leaders that partnered with us on this research project. We wish to recognize and thank the anonymous participants in our study for taking the time to respond to our survey and speak with us. All opinions expressed represent those of the authors and not necessarily the institutions with which they are affiliated or the U.S. Department of Education. All errors are solely the responsibility of the authors.

Declaration of Conflicting Interests


The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Funding for this article was provided by a grant from the U.S. Department of Education's Institute of Education Sciences (R305A160100).

ORCID iDs

Madeline Mavrogordato  <https://orcid.org/0000-0002-5401-8877>

Morgaen L. Donaldson  <https://orcid.org/0000-0003-4999-7473>

Hana Kang  <https://orcid.org/0000-0001-8681-8804>

References

- Anderson, L. M., & Turnbull, B. J. (2016). *Evaluating and supporting principals*. Wallace Foundation.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage.
- Creswell, J. W., & Guetterman, T. C. (2019). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (6th ed.). Pearson.
- Davis, S. H., & Hensley, P. A. (1999). The politics of principal evaluation. *Journal of Personnel Evaluation in Education*, 13(4), 383–403. <https://doi.org/10.1023/A:1008161417522>
- Deci, E. L., & Flaste, R. (1995). *Why we do what we do: The dynamics of personal autonomy*. GP Putnam's Sons.
- Deci, E. L., Ryan, R. M., & Williams, G. C. (1996). Need satisfaction and the self-regulation of learning. *Learning and Individual Differences*, 8(3), 165–183. [https://doi.org/10.1016/S1041-6080\(96\)90013-8](https://doi.org/10.1016/S1041-6080(96)90013-8)
- Donaldson, M., Mavrogordato, M., Dougherty, S. M., Ghanem, R. A., & Youngs, P. (2021a). Principal evaluation under the elementary and secondary every student succeeds act: A comprehensive policy review. *Education Finance and Policy*, 16(2), 347–361. https://doi.org/10.1162/edfp_a_00332

- Donaldson, M. L., Mavrogordato, M., Youngs, P., Dougherty, S., & Al Ghanem, R. (2021b). "Doing the 'real' work": How superintendents' sensemaking shapes principal evaluation policies and practices in school districts. *AERA Open*, 7(1), 1–16. <https://doi.org/10.1177/2332858420986177>
- Donaldson, M. L. (2020). *Multidisciplinary perspectives on teacher evaluation: Understanding the research and theory*. Routledge.
- Firestone, W. A. (2014). Teacher evaluation policy and conflicting theories of motivation. *Educational Researcher*, 43(2), 100–107. <https://doi.org/10.3102/0013189X14521864>
- Fuller, E. J., Hollingworth, L., & Liu, J. (2015). Evaluating state principal evaluation plans across the United States. *Journal of Research on Leadership Education*, 10(3), 164–192. <https://doi.org/10.1177/1942775115614291>
- Goff, P., Guthrie, J. E., Goldring, E., & Bickman, L. (2014). Changing principals' leadership through feedback and coaching. *Journal of Educational Administration*, 52(5), 682–704. <https://doi.org/10.1108/JEA-10-2013-0113>
- Goldring, E. B., Clark, M. A., Rubin, M., Rogers, L. K., Grissom, J. A., Gill, B., Kautz, T., McCullough, M., Neel, M., & Burnett, A. (2020). *Changing the principal supervisor role to better support principals: Evidence from the Principal Supervisor Initiative*. Mathematica.
- Goldring, E. B., Mavrogordato, M., & Haynes, K. T. (2015). Multisource principal evaluation data: Principals' orientations and reactions to teacher feedback regarding their leadership effectiveness. *Educational Administration Quarterly*, 51(4), 572–599. <https://doi.org/10.1177/0013161X14556152>
- Grissom, J. A., Blissett, R. S., & Mitani, H. (2018). Evaluating school principals: Supervisor ratings of principal practice and principal job performance. *Educational Evaluation and Policy Analysis*, 40(3), 446–472. <https://doi.org/10.3102/0162373718783883>
- Hamilton, L. S., Engberg, J., Steiner, E. D., Nelson, C. A., & Yuan, K. (2012). *Improving school leadership through support, evaluation, and incentives: The Pittsburgh Principal Incentive Program*. RAND.
- Henry, G. T., & Viano, S. L. (2015). *An evaluation of the North Carolina Educator Evaluation System for School Administrators: 2010–11 through 2013–14*. Vanderbilt University.
- Hermann, M., & Ross, C. (2016). *Measuring principals' effectiveness: Results from New Jersey's first year of statewide principal evaluation REL 2016–156*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic.
- Hill, C. E., Knox, S., Thompson, B. J., Nutt Williams, E., & Hess, S. A. (2005). Consensual qualitative research: An update. *Journal of Counseling Psychology*, 52(2), 196–205. <https://doi.org/10.1037/0022-0167.52.2.196>
- Honig, M., & Rainey, L. (2014). Central office leadership in principal professional learning communities: The practice beneath the policy. *Teachers College Record*, 116(4), 773–774.

- Honig, M. I. (2012). District central office leadership as teaching: How central office administrators support principals' development as instructional leaders. *Educational Administration Quarterly*, 48(4), 733–774. <https://doi.org/10.1177/0013161X12443258>
- Honig, M. I., & Hatch, T. C. (2004). Crafting coherence: How schools strategically manage multiple, external demands. *Educational Researcher*, 33(8), 16–30. <https://doi.org/10.3102/0013189X033008016>
- Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Field Methods*, 18(1), 3–20. <https://doi.org/10.1177/1525822X05282260>
- Jacques, C., Clifford, M., & Hornung, K. (2012). *State policies on principal evaluation: Trends in a changing landscape*. National Comprehensive Center for Teacher Quality.
- Kimball, S. M., Arrigoni, J., Clifford, M., Yoder, M., & Milanowski, A. (2015). *District leadership for effective principal evaluation and support*. U.S. Department of Education. Retrieved from <http://files.eric.ed.gov/fulltext/ED566525.pdf>.
- Kimball, S. M., Milanowski, A., & McKinney, S. A. (2009). Assessing the promise of standards-based performance evaluation for principals: Results from a randomized trial. *Leadership and Policy in Schools*, 8(3), 233–263. <https://doi.org/10.1080/15700760802416099>
- Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, 57(9), 705–717. <https://doi.org/10.1037/0003-066X.57.9.705>
- McCullough, M., Lipscomb, B., Chiang, H., & Gill, B. (2016). *Do principals' professional practice ratings reflect their contributions to student achievement? Evidence from Pennsylvania's Framework for Leadership: Working Paper 46*. Mathematica Policy Research, Inc.
- Miles, M. B., Huberman, A. M., & Saldana, J. M. (2020). *Qualitative data analysis: A methods sourcebook* (4th Edition). Sage.
- Moss, P. A., & Haertel, E. H. (2016). Engaging methodological pluralism. In D. Gitomer, & C. Bell (Eds.), *Handbook of research on teaching* (5th edition, pp. 127–247). American Educational Research Association.
- Ordóñez, L. D., Schweitzer, M. E., Galinsky, A. D., & Bazerman, M. H. (2009). Goals gone wild: The systematic side effects of overprescribing goal setting. *Academy of Management Perspectives*, 23(1), 6–16. <https://doi.org/10.5465/amp.2009.37007999>
- Reeves, D. B. (2005). *Assessing educational leaders: Evaluating performance for improved individual and organizational results*. Corwin Press.
- Rorrer, A. K., Skrla, L., & Scheurich, J. J. (2008). Districts as institutional actors in educational reform. *Educational Administration Quarterly*, 44(3), 307–357. <https://doi.org/10.1177/0013161X08318962>
- Rubin, M., Goldring, E., Neel, M. A., Rogers, L. K., & Grissom, J. A. (2021). Changing principal supervision to develop principals' instructional leadership capacity. In P. Youngs, & J. Kim, & M. Mavrogordato (Eds.), *Exploring*

- principal development and teacher outcomes: How school leaders can strengthen instruction, teacher retention, and student achievement* (pp. 41–55). Routledge.
- Strauss, A. L., & Corbin, J. M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Sage.
- Sun, M., & Youngs, P. (2009). How does district principal evaluation affect learning-centered principal leadership? Evidence from Michigan school districts. *Leadership and Policy in Schools, 8*(4), 411–445. <https://doi.org/10.1080/15700760902890490>
- Thomas, W. T., Holdaway, E. A., & Ward, K. L. (2000). Policies and practices involved in the evaluation of school principals. *Journal of Personnel Evaluation in Education, 14*(3), 215–240. <https://doi.org/10.1023/A:1008167130999>
- Yin, R. K. (2018). *Case study research and applications: Design and methods, 6th Edition*. Sage.
- Zepeda, S. J., Lanoue, P. D., Price, N. F., & Jimenez, A. M. (2014). Principal evaluation—linking individual and program-level progress: Making the connections and embracing the tensions. *School Leadership & Management, 34*(4), 324–351. <https://doi.org/10.1080/13632434.2014.928681>

Author Biographies

Madeline Mavrogordato is an associate professor of K-12 educational administration and policy in the College of Education at Michigan State University. Her research uses mixed methods to investigate how district, state, and federal policies shape educational outcomes for underserved students and how to develop, support, and evaluate effective school leaders who are prepared to serve the new demography of American public schools.

Peter Youngs is a professor in the Department of Curriculum, Instruction and Special Education at the University of Virginia. His research focuses on teacher quality, educational leadership, and school reform. Recent publications have appeared in *AERA Open*, *Journal of Teacher Education*, and *Teachers College Record*.

Morgaen Donaldson is Philip E. Austin Chair, professor of educational leadership in the Neag School of Education, and director of the Center for Education Policy Analysis, Research, and Evaluation at the University of Connecticut. Her research examines teacher and leader evaluation and development and the ways in which these and other policies shape the work of district leaders, principals, and teachers with a particular focus on underresourced districts.

Hana Kang is an assistant professor in the College of Education at the University of Colorado Colorado Springs. Her research examines educational disparities derived from differential opportunity to learn and how different school and social contexts come into play in the implementation of policy measures for school improvement.

Shaun M. Dougherty is a professor of education and policy at Boston College. His research focuses on the use of quantitative research methods to evaluate the impact of educational policies and programs. He emphasizes understanding how the requirements, incentives, and behaviors that those policies produce develop human capital.